

**nccit
2014**

**THE 10TH NATIONAL CONFERENCE ON
COMPUTING AND INFORMATION
TECHNOLOGY**

PROCEEDINGS OF NCCIT 2014

THE 10TH NATIONAL CONFERENCE ON COMPUTING AND INFORMATION TECHNOLOGY

8TH-9TH MAY 2014
ANGSANA LAGUNA PHUKET, THAILAND

WWW.NCCIT.NET

FACULTY OF INFORMATION TECHNOLOGY
KING MONGKUT'S UNIVERSITY OF TECHNOLOGY NORTH BANGKOK

บทความวิจัย

การประชุมทางวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ
ครั้งที่ 10

8-9 พฤษภาคม 2557
โรงแรมอังกูนา ลากูนา ภูเก็ต



คณะเทคโนโลยีสารสนเทศ

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

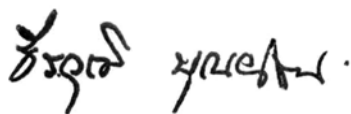
TABLE OF CONTENTS

Message from KMUTNB President	ii
Message from General Chair	iv
Conference Organizers	viii
Conference Organization Committee.....	x
Steering Committee.....	xi
Technical Program Committee.....	xii
Technical Program Contents.....	xiii
Keynote Speakers	xxxiv
Regular Papers.....	1
Author Index	1036

สารจากอธิการบดี

เป็นที่ขอมรับกันโดยทั่วไปว่า การพัฒนาคุณภาพของประชากรในชาติได้กลายเป็นเงื่อนไขสำคัญ ต่อพัฒนาการของประเทศในปัจจุบันนี้ ท่ามกลางการเปลี่ยนแปลงทางด้านเทคโนโลยี สภาพเศรษฐกิจ ระหว่างประเทศที่มีการแข่งขันสูง สภาพปัญหาสังคมวัฒนธรรม และปัญหาสิ่งแวดล้อมของโลก การพัฒนา ทางด้านกำลังคนของประเทศที่เข้มแข็งทั้งทางร่างกายจิตใจและสติปัญญา เพื่อพร้อมที่จะเผชิญกับปัญหา และความเปลี่ยนแปลงอย่างมีวิจรรย์ญาณ ไม่ว่าจะในสังคมเมืองไปจนถึงสังคมระดับรากหญ้า จะเป็น หลักประกันความมั่นคงโดยรวมของประเทศในระยะยาว โดยเฉพาะในส่วนของประชาคมเศรษฐกิจ อาเซียน หรือ AEC ที่กำลังจะเกิดขึ้นในปี พ.ศ. 2558 ในกรณีนี้เทคโนโลยีสารสนเทศนับว่าเป็นกลไกสำคัญ ประการหนึ่ง ในการสนับสนุนให้มีการจัดการทางการศึกษาที่ดีและมีคุณภาพไปสู่ความเป็นสากล ดังจะ เห็นได้จากทุกประเทศที่เจริญแล้ว ต่างมีการลงทุนทรัพยากรเพื่อการวางโครงสร้างและปัจจัยพื้นฐานด้าน เทคโนโลยีสารสนเทศ เพื่อการศึกษาของประชาชนในชาติพร้อมพัฒนานวัตกรรมการเรียนการสอน ที่ มุ่งเน้นการนำเทคโนโลยีสารสนเทศมาใช้เป็นประโยชน์ได้อย่างเป็นรูปธรรมในการจัดการศึกษาทุกระดับ จุดมุ่งหมายสำคัญก็เพื่อยกระดับคุณภาพของประชากรในชาติให้ก้าวไกล ทัวถึง และเท่าเทียมกัน และใน ประเทศไทยก็เช่นเดียวกัน การพัฒนาศักยภาพด้านการวิจัยด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศจึง เป็นปัจจัยสำคัญ ที่จะนำพาประเทศไปสู่การแข่งขันในภูมิภาคและสากลต่อไป

ขอขอบคุณหน่วยงาน คณะกรรมการจัดงาน และคณะบุคคลที่มีส่วนสนับสนุนในการจัดการ ประชุมทางวิชาการครั้งนี้ ขออวยพรให้การประชุมทางวิชาการครั้งนี้ประสบความสำเร็จบรรลุวัตถุประสงค์ ทุกประการ



(ศาสตราจารย์ ดร.ธีรวิทย์ นุญยโสภณ)

อธิการบดี มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

MESSAGE FROM KMUTNB PRESIDENT

Nowadays, it is generally accepted that the development of a nation stems from technical advancements and this has become the main key factor which dictates the development of any country. Many issues can affect development such as international economics, highly competitive markets, social and cultural differences, and global environmental problems. Thus, strengthening a country's capability to gain knowledge both physically and mentally provides the ability to deal with any critical issues, not only in an urban society but also surrounding countryside societies, too. To ensure the overall stability of a country in the long term, technology is considered one of the most important mechanisms to support the management of educational quality and to provide the ability to continually improve it. This is especially the case with the ASEAN Economic Community or AEC being formed in 2015. It can be seen from many civilized countries that have invested resources and fundamental infrastructures for the deployment of Information Technology to their education system. The development of innovative teaching and learning that focuses on bringing Information Technology to the forefront of education benefits the population as a whole. The main goal is to improve the quality of life and progress evenly and equally to a better future. Also in Thailand, the development of the research on computing and information technology is a very important factor that will help the country become competitive on both regional and international stages.

I would like to say a special thank you to everybody involved in this conference, from partners to stakeholders, without you this would not be possible. And I hope this conference provides a good opportunity for all your voices to be heard.



(Professor Dr. Teravuti Boonyasopon)
President, King Mongkut's University of Technology North Bangkok

สารจากประธานจัดงาน

หนังสือเล่มนี้เป็นเอกสารรวมเล่มรายงานผลวิจัย จากการประชุมวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 10 (NCCIT 2014) ซึ่งจัดขึ้นที่อัสสนาลากูน่าภูเก็ต จังหวัดภูเก็ต เมื่อวันที่ 8 - 9 พฤษภาคม 2557 ซึ่งจัดร่วมกับประชุมวิชาการระดับนานาชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 10 (The 10th International Conference on Computing and Information Technology: IC²IT 2014) เป็นการจับคู่กันของมหาวิทยาลัยเครือข่ายทั้งในประเทศและต่างประเทศ

สำหรับปีนี้ NCCIT 2014 ได้รับบทความวิจัยจำนวน 249 ฉบับ จาก 76 สถาบันทั่วประเทศ โดยถูกส่งให้ผู้ทรงคุณวุฒิพิจารณาจำนวน 3 ท่าน และบทความที่ได้รับการตอบรับจากผู้ทรงคุณวุฒิอย่างน้อย 2 ใน 3 ท่าน ผลงานวิจัยที่ได้รับการตอบรับมีจำนวน 165 ผลงานคิดเป็นร้อยละ 66 ที่นำเสนอปกเกล้าและบรรจุในเอกสารรวมเล่มรายงานผลวิจัยงานประชุมวิชาการ NCCIT2014 โดยผลงานวิจัยที่มีคุณภาพดีเยี่ยมจำนวนหนึ่งจะถูกคัดเลือกเพื่อพิจารณาลงพิมพ์ในวารสารเทคโนโลยีสารสนเทศ (ISSN 1685-8573) ซึ่งเป็นวารสารวิชาการที่อยู่ในฐานข้อมูล Thailand Journal Citation Index (TCI) กลุ่มที่ 1 ในปี 2555 มีค่า Impact Factor เท่ากับ 0.024

คณะกรรมการจัดงานประชุม NCCIT2014 ขอขอบคุณนักวิจัยทุกท่านที่ส่งบทความวิจัยเพื่อพิจารณา ทั้งผู้ที่ได้รับการตอบรับและผู้ที่ไม่ได้รับการตอบรับในการนำเสนอบทความวิจัย ขอขอบคุณคณะกรรมการพิจารณาบทความที่เสียเวลาอันมีค่าเพื่อพิจารณาบทความอย่างเข้มข้นได้มาตรฐาน ทั้งนี้เพื่อให้ได้ผลงานวิจัยที่มีคุณภาพในการลงพิมพ์ในเอกสารรวมเล่มรายงานผลวิจัย และการนำเสนอในที่ประชุมวิชาการ NCCIT 2014

ขอขอบพระคุณ ศาสตราจารย์ ดร.ธีรวิทย์ บุญยโสภณ อธิการบดีมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ สำหรับการสนับสนุนการประชุมวิชาการ NCCIT มาอย่างต่อเนื่องโดยตลอดนับเป็นเวลา 10 ปี

ขอขอบคุณมหาวิทยาลัยเครือข่ายในประเทศไทยทั้งสิบแห่งที่ให้การสนับสนุนเป็นอย่างดีทั้งด้านการเงิน และบุคลากรในการร่วมเป็นกรรมการเครือข่าย กรรมการดำเนินงาน และกรรมการพิจารณาบทความ ได้แก่ มหาวิทยาลัยมหาสารคาม มหาวิทยาลัยอุบลราชธานี มหาวิทยาลัยราชภัฏกาญจนบุรี มหาวิทยาลัยราชภัฏนครปฐม มหาวิทยาลัยราชภัฏมหาสารคาม มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา มหาวิทยาลัยเทคโนโลยีราชมงคลกรุงเทพ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี มหาวิทยาลัยสงขลานครินทร์วิทยาเขตภูเก็ต และสถาบันบัณฑิตพัฒนบริหารศาสตร์

ขอขอบคุณเป็นพิเศษสำหรับผู้ช่วยศาสตราจารย์ ดร.ศิริปัฐ บุญครอง นางสาววัชรวิมล จิตต์สกุล และทีมงานวิจัย ที่ได้ประสานงานกันนักวิจัยและผู้ที่เกี่ยวข้องเพื่อจัดรวบรวม ทำเล่มรายงานผลงานวิจัยสำหรับงานประชุมให้สำเร็จในเวลาที่เหมาะสม และขอบคุณอาจารย์กาญจนา วิริยะพันธ์ อาจารย์จิระศักดิ์ นำประดิษฐ์ และนายอมรศักดิ์ อมรชนานันท์ ซึ่งเป็นผู้พัฒนาและดูแลเว็บของงานประชุมเป็นอย่างดี

ขอขอบคุณคณาจารย์ และเจ้าหน้าที่ของคณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ และมหาวิทยาลัยเครือข่ายที่ให้การสนับสนุนด้านบุคลากรและเงินสนับสนุน

หวังว่า NCCIT จะได้ให้โอกาสที่ดีสำหรับคณาจารย์ นักวิจัย นักวิชาการ และนักศึกษา ได้นำเสนอผลงานเป็นการพัฒนาตนเองสู่ความเป็นนักวิจัยที่เข้มแข็งยิ่ง ๆ ขึ้นไป NCCIT จะเป็นแพลตฟอร์มสำหรับการแลกเปลี่ยนความรู้ ในด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ และจะสร้างแรงบันดาลใจให้นักวิจัย เพื่อสร้างความคิดใหม่และการค้นพบเทคโนโลยีใหม่ ๆ และพบพันธมิตร เพื่อสร้างเครือข่ายการวิจัยสำหรับการทำงานร่วมกันในอนาคตต่อไป



รองศาสตราจารย์ ดร. พยุง มีถ้ำ

คณบดีคณะเทคโนโลยีสารสนเทศ มจพ.

ประธานจัดงานประชุมวิชาการ NCCIT2014

MESSAGE FROM GENERAL CHAIR

This volume contains the papers of the 10th National Conference on Computing and Information Technology (NCCIT2014) held at Angsana Laguna Phuket, Thailand, on May 8th-9th, 2014. Traditionally, it is organized in conjunction with the International Conference on Computing and Information Technology (IC²IT), which is supported by many universities both in Thailand and overseas.

This year, the program committee of NCCIT2014 received 249 submissions from authors of over 76 institutions. Each submission was reviewed by at least 2, mostly 3 members of the program committee to avoid contradictory results. On these judgments, the committee decided to accept 165 papers (66% accepted rate) for oral presentation and inclusion in the conference proceedings of NCCIT2014. The papers from NCCIT2014 with excellent quality will be selected for consideration and hopefully published in the Information Technology Journal (ISSN 1685-8573), which is included in a database of academic journals in Thailand Journal Citation Index (TCI). The Information Technology Journal of 2013 had an Impact Factor at 0.024.

A special thanks is given to Professor Dr. Teeravuti Boonyasopon, KMUTNB President, for his support of NCCIT from the first year onward, and for providing us with the necessary resources from KMUTNB.

I would like to also thank the 10 co-operative universities for supporting both in finance and personnel, which are Mahasarakham University, Ubon Ratchathani University, Kanchanaburi Rajabhat University, Nakhon Pathom Rajabhat University, Mahasarakham Rajabhat University, Rajamangala University of Technology Lanna, Rajamangala University of Technology Krungthep, Rajamangala University of Technology Thanyaburi, Prince of Songkla University Phuket Campus, and National Institute of Development Administration.

A very special and warm thank you also goes to Assistant Professor Dr. Sirapat Boonkrong, Ms. Watchareewan Jitsakul, and the supportive research team; the proceedings could not have been completed in a timely manner and correct format, without the painstaking work. Moreover, a special thank is given to our web masters: Ms. Kanchana Viriyapant, Mr. Jeerasak Numpradit, and Mr. Armornsak Armornthananun. A lot of technical and organizational work has been done by the staff of the Information Technology Faculty at KMUTNB. We would like to also thank all authors for their submissions and the members of the program committee for their great work and valuable time.

NCCIT delivers an excellent opportunity for lectures, researchers, scholars, and students to present their work. University lecturers can contribute their research works in order to promote their academic titles of professorships. It provides an opportunity for graduate students to present their work and fulfill the requirements of graduation. Moreover, it is a forum for doctoral students to improve themselves by presenting and strengthening their research works before publishing in an academic journal. Hopefully, NCCIT strives to be a platform in which exchange of knowledge in the field of computers and information technology is forefront, aiming to inspire researchers with new ideas and possible future collaboration.



Associate Professor Dr. Phayung Meesad
Dean, Faculty of Information Technology
King Mongkut's University of Technology North Bangkok
NCCIT2014 General Chair

CONFERENCE ORGANIZERS



King Mongkut's University of Technology North Bangkok, Thailand
Faculty of Information Technology



Chemnitz University, Germany

FernUniversität in Hagen, Germany



Edith Cowan University, Australia

Oklahoma State University, USA



Hanoi National University of Education, Vietnam



Gesellschaft für Informatik (GI), Germany

Maharakham University, Thailand
Faculty of Information Technology



Ubon Ratchathani University, Thailand
Mathematics Science and Computer

Kanchanaburi Rajabhat University, Thailand



Nakhon Pathom Rajabhat University, Thailand
The Office of Academic Promotion and Registration

Rajabhat Maharakham University, Thailand
Faculty of Information Technology



Rajamangala University of Technology Lanna, Thailand
Faculty of Engineering

Rajamangala University of Technology Krungthep, Thailand
Center of Academic Resources and Information Technology



CONFERENCE ORGANIZERS



Rajamangala University of Technology Thanyaburi, Thailand

Prince Of Songkla University, Phuket Campus, Thailand



National Institute of Development Administration, Thailand

Department of Information Technology and Systems Management, School of Applied Statistics

Council of IT Deans of Thailand (CITT) , Thailand



KMUTNB 55th Anniversary

A University of Creative Invention to Innovation

CONFERENCE ORGANIZATION COMMITTEE

General Chair :	Assoc. Prof. Dr. Phayung Meesad, KMUTNB
Technical Program and Secretary :	Asst. Prof. Dr. Sirapat Boonkrong, KMUTNB
Conference Treasurer :	Asst. Prof. Dr. Nattavee Utakrit, KMUTNB

STEERING COMMITTEE

Assoc. Prof. Dr. Arunee Intrapairot, RMUTT

Asst. Prof. Dr. Janpen Intaraprasert, UBU

Asst. Prof. Dr. Jugkarin Sukmok, NIDA

Asst. Prof. Dr. Rattana Wetprasit, PSU

Asst. Prof. Dr. Sirapat Boonkrong, KMUTNB

Asst. Prof. Dr. Sujin Butdisuwan, MSU

Asst. Prof. Dr. Tharat Arreerard, MRU

Asst. Prof. Dr. Uraiwan Inyaem, RMUTT

Asst. Prof. Dr. Worapapha Arreerard, MRU

Asst. Prof. Photjanee Sukchovna, KRU

Asst. Prof. Surasak Yousawat , RMUTL

Dr. Kairung Hengprapohm, NPRU

Dr. Siwiga Dusadenoad, NIDA

Dr. Supoj Hengprapohm, NPRU

Dr. Thawatchai Sarawong, RMUTK

Dr. Wongot Sriurai, UBU

TECHNICAL PROGRAM COMMITTEE

Alisa Kongthon, NECTEC	Paranya Ammaruekarat, NPRU
Anirach Mingkhuang, KMUTNB	Parinya Sa-Nguansat, PIM
Apida Roonwat, MSU	Passakorn Prathombutr, NECTEC
Banatus Soiraya, UBS Co.,Ltd.	Patharawut Saengsiri, TISTR
Burin Rujjanapan, Nation	Phattanapong Chompoowises, MSU
Chatchawin Namman, UBU	Phayung Meesad, KMUTNB
Chatklaw Charoenpol, MSU	Pilapan Phonarin, RMUTK
Chanattha Thongsuk, MU	Pongpisit Wuttidittachotti, KMUTNB
Chayakorn Netramai, KMUTNB	Poolsak Koseeyaporn, KMUTNB
Choochart Haruechaiyasa, NECTEC	Prachyanun Nilsook, KMUTNB
Dech Thammasiri, NPRU	Prasong Praneetpolgrang, SPU
Gamkarn Somprasertsri, MSU	Pudsadee Boonrawd, KMUTNB
Jamornkul Laokietkul, CRU	Punnee Sittidech, NU
Jaran Sanrach, KMUTNB	Sakchai Tangwannawit, KMUTNB
Jaree Thongkam, MSU	Saowalak Arampongsanuwat, SU
Jeerasak Nampradit, KMUTNB	Sirapat Boonkrong, KMUTNB
Jirapan Srisompan, KMUTNB	Siriporn Supratit, RSU
Jiratta Phuboon-ob, MSU	Sithiporn Kerdsun-ang, KMUTNB
Kairung Hengpraprom, NPRU	Somchai Prakanchaen, KMUTNB
Kanchana Viriyapant, KMUTNB	Somjit Arch-int, KKU
Kantpong Woraratpanya, KMITL	Somnuk Puangpronpitag, MSU
Kiattisak Yochanang, KMUTNB	Sompong Valuvanathorn, UBU
Kittiwan Nimkerdphol, RMUTT	Somsak Akatimagool, KMUTNB
Kriengsak Treeprapin, UBU	Soradech Krootjohn, KMUTNB
Mahasak Ketcham, KMUTNB	Sucha Smachat, KMUTNB
Maleerat Sodanil, KMUNB	Suksaeng Kukanok, IBA Co.,Ltd.
Manasawee Kaenampornpan, MSU	Sumitra Nuanmesri, SSRU
Manoon Srivirat, UBU	Sunantha Sodsee, KMUTNB
Meechai Lohakan, KMUTNB	Supot Nitsuwat, KMUTNB
Montean Rattanasiriwongwut, KMUNB	Tharat Arreerard, MRU
Nadh Ditchaen, UBU	Thawatchai Sahapong, MRU
Nalinpat Porrawatpreyakorn, KMUNB	Thawatchai Sarawong, RMUTK
Narong Photi, MOE	Thippaya Chintakovid, KMUTNB
Nattavee Utakrit, KMUTNB	Thirapon Wongsaaardsakul, BU
Nawaporn Wisitpongphan, KMUNB	Tong Srikacha, TOT
Nidapan Sureerattanan, TNI	Tossaporn Joochim, UBU
Nivet Chirawichitchai, SPU	Udom Jeenpradub, KMUTNB
Noppasak Tantisattayanon, RMUTR	Uraivan Inyaem, RMUTT
Nucharee Premchaiswadi, DPU	Vatinee Nuipian, KMUTNB
Padej Phomsakha-Na-Sakonnakorn, CU	Veerawan Janthanasub, RMUTK
Pairoj Likitthanasate, TOT	Watchareewan Jitsakul, KMUTNB
Panana Tangwannawit, PCRU	Winai Bhosuwan, KU
Panida Songram, MSU	Wongot Sriurai, UBU
Panomkhawn Riyamongkol, NU	Worapapha Arreerard, MRU

TECHNICAL PROGRAM CONTENTS

Thursday May 8th, 2014	
8:00-9:00	Registration
9:00-9:10	Report (in Thai) by Assoc.Prof.Dr. Phayung Meesad , <i>Dean of Faculty of Information Technology, King Mongkut's University of Technology North Bangkok</i> Report by Prof.Dr.Herwig Unger , <i>IC²IT 2014 Technical Program Chair, Fern University in Hagen, Germany</i>
9:10-9:40	Opening Ceremony by Prof. Dr. Teravuti Boonyasopon , <i>President of King Mongkut's University of Technology North Bangkok</i>
9:40-10:40	Invited Keynote Speech by Prof. Dr.Peter Kropf <ul style="list-style-type: none"> • <i>Dean of Faculty of Sciences, University of Neuchâtel, Switzerland</i> Topic: Wireless Mesh Networks and Cloud Computing for Real Time Environmental Simulations
10:40-11:00	<i>Coffee Break</i>
11:00-12:00	Invited Keynote Speech by Assoc. Prof. Dr.Hiroya Takamura <ul style="list-style-type: none"> • <i>Advanced Information Processing Division, Precision and Intelligence Laboratory, Tokyo Institute of Technology</i> Topic: Mathematical Models for Text Summarization
12:00-13:00	<i>Lunch</i>
13:00-17:30	Parallel Session Presentation
17:30-22:00	<i>Welcome Dinner</i>

Thursday May 8 th , 2014		
NCCIT2014 Room I		
Classification and Clustering Applications		
Time	Title/Author	Page
13:00-13:15 NCCIT2014-254	Decision Tree Classification of Imbalanced Data Sets Using Different Weight On Synthesized Data <i>Suraphong Cheawsakunwattana and Sukree Sinthupinyo</i>	1
13:15-13:30 NCCIT2014-247	Rice Yield Prediction Based on Artificial Neural Networks <i>Attapol Kunlerd and Pusadee Seresangtakul</i>	7
13:30-13:45 NCCIT2014-156	Genetic Algorithm for Weighted K-Means <i>Natthawoot Punroob, Athaporn Gingboo and Guangqian Chen</i>	13
13:45-14:00 NCCIT2014-123	Classification System Para Rubber Varieties using Naïve Bayes <i>Worraphot Sumethawatthanaphong, Nutthavipa Songsook and Wannasiri Thurachon</i>	20
14:00-14:15 NCCIT2014-206	A Model for Swine Disease Analysis using Neural Network <i>Jutarat Tangkittiwat and Nalinpat Porrawatpreyakorn</i>	26
14:15-14:30 NCCIT2014-245	The Comparison of Clustering Methods Used for Food Clustering Analysis <i>Sumalee Eisariyodom and Ureerat Suksawatchon</i>	32
14:30-14:45 NCCIT2014-262	The Analysis of Relationship-based Clustering in e-Distance Mentoring System <i>Photjane Sukchaona and Monchai Tiantong</i>	39
14:45-15:00 NCCIT2014-252	A Document Binarization Method Based on Local Hierarchical Clustering <i>Amorntape Yamkasorn and Anupan Rodtook</i>	45
15.00-15.30	Coffee Break	
15:30-15:45 NCCIT2014-26	News Timeline Visualization Tool Based on News Category Classification <i>Vitthawat Sajjacharoenpong and Twittie Senivongse</i>	51
15:45-16:00 NCCIT2014-193	Test Case Generation Using Genetic Algorithms <i>Nuntanee Chuaychoo and Supaporn Kansomkeat</i>	58
16:00-16:15 NCCIT2014-63	The Application Q-Gram Algorithm and Splay-Tree Data Structure to Enhance Electronic Dictionary of TCM Terminology <i>Peerapad Pandison, Woramet Withunwet, Aumpai Chobdean and Ekapong Chuasuwan</i>	64
16:15-16:30 NCCIT2014-185	A Prototype of Offline Tolerant Web using HTML5 <i>Pamok Rattanatriyapibal and Maleerat Sodanil</i>	70
16:30-16:45 NCCIT2014-256	A Development of a Non-life Insurance News Aggregator Prototype <i>Jatupit Aiyarakom and Twittie Senivongse</i>	76

Thursday May 8th, 2014		
NCCIT2014 Room I		
Classification and Clustering Applications		
Time	Title/Author	Page
16.45-17.00 NCCIT2014-148	Automated Generation of Test Cases System by using Classification Tree <i>Suchada Ratanakongnate and Kittichai Sopitwetmontree</i>	83
17.00-17.15 NCCIT2014-43	Tree Structure for Association Identification between Module <i>Zagon Budsabong and Pichayotai Mahathanapawat</i>	89
17:30-22:00	<i>Welcome Dinner and Best Paper Award Presentation</i>	

Thursday May 8 th , 2014		
NCCIT2014 Room II		
Web Service, Semantic Web, and Web Application		
Time	Title/Author	Page
13:00-13:15 NCCIT2014-240	An Application of Hotel Searching Based on Opinion Mining <i>Janasthar Srisuan and Anantaporn Hunskunatai</i>	95
13:15-13:30 NCCIT2014-69	Question Classification for Answer Searching Using Semantic Web and Data Mining <i>Supote Bualerng and Wararat Songpan</i>	101
13:30-13:45 NCCIT2014-130	Automatic Tourist Attraction Categorization System using Thai Description via Web Service <i>Choopan Rattanapoka and Mathawee Sutikun</i>	109
13:45-14:00 NCCIT2014-91	The Development of Next Generation Sequencing Tool for Annotation and Variant Filtrating of Human Nucleotide Alteration <i>Kittaboon Panjarattankorn, Ubolwan Chaovanakij and Kitsuchart Pasupa</i>	115
14:00-14:15 NCCIT2014-246	Surveying and Mapping Network Availability using Crowdsourcing <i>Thaphat Mupattararot and Veera Muangsin</i>	121
14:15-14:30 NCCIT2014-37	Accident Warning System and Specifying the Location on a Map using Mobile Device for the Emergency Medical Services <i>Tawatchai Ruangtananurak, Ngamnij Arch-Int, Somjit Arch-Int and Saiyan Saiyod</i>	127
14:30-14:45 NCCIT2014-141	An Approach to Testing RESTful Web Services in Agile Development <i>Sarita Khuptayanon and Yachai Limpiyakorn</i>	133
14:45-15:00 NCCIT2014-182	A Model-Driven Development of Web-Based Applications on Google App Engine Platform <i>Somrudee Kaewkao and Twittie Senivongse</i>	140
15.00-15.30	Coffee Break	
15:30-15:45 NCCIT2014-72	Automatic System of Post Master Supported Affiliate Marketer by Number of Post Method <i>Kittinan Noimanee and Natchasit Chukiathajorn</i>	146
15:45-16:00 NCCIT2014-202	A Model for Ranking Search Results in a Research Paper Search Engine Using Bibliographic Information <i>Khwanruan So-Ubol and Worasit Choochaiwattana</i>	150
16:00-16:15 NCCIT2014-216	A Mechanism for Securing WSDL Web Service <i>Janejira Wanglee and Ladda Preechaveerakul</i>	156
16:15-16:30 NCCIT2014-155	Load Balancing Architecture for Web Servers with NGINX <i>Surachat Boonsaen and Nuengwong Tuaycharoen</i>	162
16:30-16:45 NCCIT2014-179	Comparison of Feature Selection Methods for Inappropriate Webpage Classification by Data Mining Technique <i>Namthip Marknakorn and Maleerat Sodanil</i>	168
16:45-17:00 NCCIT2014-231	Guidelines for Skill Development Model of Computing Design Process by Using 2.0 Web Technology and Mentoring Technique <i>Sununta Klintawon, Netchanok Jansawang and Worapapha Arreerard</i>	174
17:30-22:00	Welcome Dinner and Best Paper Award Presentation	

Thursday May 8 th , 2014		
NCCIT2014 Room III		
Wireless Networks		
Time	Title/Author	Page
13:00-13:15 NCCIT2014-77	Gateway Discovery Mechanism with Low Message Exchange Overhead for Mobile Ad Hoc Networks <i>Sumet Prabhavat, Sarach Sriswadpong, Tanong Chuvanitchanon and Worrawat Narongkhachavana</i>	180
13:15-13:30 NCCIT2014-139	Multipath Load-balancing Improvement on Mobile Ad-hoc Network <i>Supachoke Pongquan, Worrawat Narongkhachavana and Sumet Prabhavat</i>	186
13:30-13:45 NCCIT2014-147	Low information feedback and Makeup constellation method OFDM system over Multipath fading <i>Chaiyaporn Khemapatapan and Rushapon Piaboran</i>	192
13.45-14.00 NCCIT2014-86	Performance and Quality Evaluation of multimedia applications over Wireless LAN <i>Parkpoom Pornpratanwech and Panwit Tuwanut</i>	198
14.00-14.15 NCCIT2014-145	Multiple Gateways System for Wireless Sensor Network IEEE802.15.4 <i>Chanadej Yamcleee and Chaiyaporn Khemapatapan</i>	204
14.15-14.30 NCCIT2014-34	A Study of Factors Effecting Wireless Networking in any Environments <i>Surapon Kraijak and Panwit Tuwanuti</i>	210
Security		
14.30-14.45 NCCIT2014-209	ARP Poisoning Detection by MAC Address Gateway's Hash value <i>Hatsada Phruekthachanbodi and Sirapat Boonkrong</i>	217
14.45-15.00 NCCIT2014-239	The Effective of Wireless Jamming: Pulsed Jamming and Swept Jamming for Wireless Ad hoc Network <i>Aekkarat Lorphichian and Sirapat Boonkrong</i>	223
15.00-15.30	<i>Coffee Break</i>	
15:30-15:45 NCCIT2014-208	SQL Injection Attack Prevention by Symmetric Cryptography for Intranet Application <i>Chutima Ampairat and Sirapat boonkrong</i>	229
15.45-16:00 NCCIT2014-251	A Study on the Relationship Between the Bandwidth and the Number of DDoS attack nodes with NS2 <i>Pradit Pitaksathienkul and Sirapat Boonkrong</i>	235
16:00-16:15 NCCIT2014-198	Secure Password Storing using Dynamic Salt Selection with Hash Function <i>Chaowalit Somboonpattanakit and Sirapat Boonkrong</i>	240
16.15-16.30 NCCIT2014-12	Prevention of Fake Authentication on Network System by UUID and MAC Address Audit <i>Sangavood Pota and Chatchawin Namman</i>	246
16.30-16.45 NCCIT2014-132	The Analysis of Street Networks by Graph Theory: The Case Study of Street in The Nonthaburi Province <i>Aongart Aun-A-Nan, Worawut Yimyam and Sunantha Sodsee</i>	252
17:30-22:00	<i>Welcome Dinner and Best Paper Award Presentation</i>	

Thursday May 8th, 2014		
NCCIT2014 Room IV		
Information Technology		
Time	Title/Author	Page
13:00-13:15 NCCIT2014-165	The Development of Automatic Programming Exercise Verification System <i>Sarawuth Rungcharoenkit and Kitsuchart Pasupa</i>	258
13:15-13:30 NCCIT2014-79	Voice Enabled Weight and Time Apparatuses on Practice for the Northern School for the Blind under the Patronage of the Queen <i>Upady Hatthasin, Siripong Matame and Sarawit Timim</i>	264
13:30-13:45 NCCIT2014-121	A Monitoring System for School Flood Victims Using the Android Smart Phone <i>Mayuree Jeeramard and Sakchai Tangwannawit</i>	270
13.45-14.00 NCCIT2014-114	Portable Guide System for Traveling <i>Khwanchai Suksaen, Manassanant Chantho and Nattawut Moonsart</i>	276
14.00-14.15 NCCIT2014-186	Analysis of Data Processing Performance between Relational and Non-Relational Databases of Documents <i>Pragaimas Srisuktaksin and Pudsadee Boonrawd</i>	281
14.15-14.30 NCCIT2014-200	Development of Dog's Health Care Application on Mobile Phone <i>Nattavadee Hongboonmee, Wirinya Songcheang and Usa Choothong</i>	287
14.30-14.45 NCCIT2014-211	Welfare Information System : Case Study of Electricity Generating Authority of Thailand <i>Somrudee Jaroenwaree and Maleerat Sodanil</i>	293
14.45-15.00 NCCIT2014-11	A Development of Peer Visit System for Self Assessment by SECI Model : A Case Study of Nation University <i>Burin Rujjanapan</i>	299
15.00-15.30	<i>Coffee Break</i>	
15:30-15:45 NCCIT2014-88	Development of AI BOTs for the Fighting Game Competition <i>Angkana Suwanjatuporn and Paranee Apiromsanee</i>	305
15.45-16:00 NCCIT2014-87	A Comparison of Reliability in Measuring Spinal Curvature <i>Apichaya Kiartubolpaiboon and Waranyu Wongseeree</i>	310
16:00-16:15 NCCIT2014-2	Mobile Device Augmented Reality Tour Application for Ramayana Story on Wat Prakaew Wall <i>Goragod Pangsomboon, Watcharapat Laonawatthana, Prawit Oungcharoon, Sapa Chanyachatchawan, Ratthasilp Ranokphanuwat and Nuengwong Tuaycharoen</i>	316
16.15-16.30 NCCIT2014-162	The Development of Computer Programs using GUI-SCILAB in a Study to Create the Image Animation in the Subject of Digital Image Processing for Electrical Engineering Students <i>Kitti Surpare and Meechai Lohakan</i>	322
16.30-16.45 NCCIT2014-119	Development of Electronic Meeting System for the Department of Health, Ministry of Public <i>Phongkorn Aksonnittrakul and Sakchai Tangwannawit</i>	328

Thursday May 8th, 2014		
NCCIT2014 Room IV		
Information Technology		
Time	Title/Author	Page
16.45-17.00 NCCIT2014-196	Development of Business Intelligence System for Bureau of Food <i>Sutheera Phannack and Maleerat Sodanil</i>	334
17.00-17.15 NCCIT2014-270	The Design of Targets and Performance Indicators for Payment Service: Telecommunication Organization Case Study <i>Veeraporn Siddoo and Rattana Wetprasit</i>	340
17:30-22:00	<i>Welcome Dinner and Best Paper Award Presentation</i>	

Thursday May 8 th , 2014		
NCCIT2014 Room V		
Computer Education		
Time	Title/Author	Page
13:00-13:15 NCCIT2014-30	Application of eDLTV media for learning to develop game application on Tablet <i>Peerachai Chinporn and Worapapha Arreerard</i>	346
13:15-13:30 NCCIT2014-127	The Synthetic Camp Activities Model for eDLTV Learning for Small Schools Network of Rajabhat Maha Sarakham University <i>Apichat Lagdee, Somjet Poo Sri and Worapapha Arreerard</i>	352
13:30-13:45 NCCIT2014-140	A Study on the Innovation in eDLTV Media for Instruction <i>Worapapha Arreerard, Laongtip Mathurasa, Monchai Tiantong and Dusanee Supawantanakul</i>	358
13.45-14.00 NCCIT2014-38	Application of eDLTV Media for Learning to Develop Tablet Application on Science <i>Siitiporn Ittasombut and Worapapha Arreerard</i>	365
14.00-14.15 NCCIT2014-153	A Study of Green IT Framework for Rajabhat University <i>Tharach Arreerard, Laongtip Mathurasa, Monchai Tiantong and Dusanee Supawantanakul</i>	372
14.15-14.30 NCCIT2014-19	The Application of eDLTV Media for Learning to Develop on Geometry <i>Duangnet Yotkamlue and Worapapha Arreerard</i>	378
14.30-14.45 NCCIT2014-244	Results of the Synthesis of Learning Application Development with Tablet for Teachers by Using Think-Pair-Share Technique <i>Winyoo Uttara, Worapapha Arreerard and Surat Duangchatom</i>	384
14.45-15.00 NCCIT2014-13	Developing Tendency For Learning on Op-Amp and Linear IC by Appropriated Tele-communication Technology and Information Technology <i>Pitipark Pinrod and Prachyanun Nilsook</i>	390
15.00-15.30	Coffee Break	
15:30-15:45 NCCIT2014-74	Learning Management Information System <i>Panwit Tuwanut, Saksit Sawang and Nutkamol Sripum</i>	396
15.45-16:00 NCCIT2014-97	The Development of the LMS Online for Teacher and Student in Nakhon Ratchasima Rajabhat University <i>Usanad Ua-Apisitwong and Sarayut Neankratok</i>	401
16:00-16:15 NCCIT2014-99	The Examination Timetable Arrangement by Bin Packing: A Case Study of Faculty of Science, Khon Kaen University <i>Jaroonluk Suporn, Sirapat Chiewchanwattana and Khamron Sunat</i>	407
16.15-16.30 NCCIT2014-112	The Development of Learning Media on the Topic of Land Traffic on Android Operating System <i>Nutthreeya Songyai and Wongkot Sriurai</i>	413
16.30-16.45 NCCIT2014-236	Development of Learning Media in Topics of Cell and Chromosome using Augmented Reality Technology <i>Nadh Ditcharoen, Thitikorn Prakrongyat, Nonlapan Praloppan and Supaporn Porntra</i>	419

Thursday May 8th, 2014		
NCCIT2014 Room V		
Computer Education		
Time	Title/Author	Page
16.45-17.00 NCCIT2014-113	The Development of Interactive Multimedia on the Topic of Science Experimenton on Android Operating System for Grade 2 Students <i>Atiporn Panpum and Wongkot Sriurai</i>	425
17.00-17.15 NCCIT2014-243	Practical Guidelines for eDLTV Applications for Learning with Tablet Computer by Using Mentoring Technique <i>Narakorn Srivapee, Worapapha Arreerard and Surat Duangchatom</i>	431
17.15-17.30 NCCIT2014-205	A Model of Camping Activities for Learning Improvement through Rajabhat Maha Sarakham University eDLTV Networks of Schools Participating in the Information Technology Project under the Initiatives of Her Royal Highness Princess Maha Chakri Sirindhorn <i>Thawatchai Sahapong, Worapapha Arreerard, Tharach Arreerard and Thanate Yuensook</i>	437
17.30-17.45 NCCIT2014-203	Empower the Community with IT for the Sustainable Sufficient Agriculture <i>Nilubon Tongchai</i>	444
17:45-22:00	<i>Welcome Dinner and Best Paper Award Presentation</i>	

Friday May 9th, 2014	
8:00-9:00	Registration
9:00-12:00	Parallel Session Presentation
12:00-13:00	<i>Lunch</i>
13:00-16:40	Parallel Session Presentation

Friday May 9th, 2014		
NCCIT2014 Room I		
Machine Learning in Prediction		
Time	Title/Author	Page
09:00-09:15 NCCIT2014-269	Interval Type-2 Fuzzy Systems Trained by Hybrid Heuristic Algorithm for Stock Exchange of Thailand <i>Adisak Sangsongfa and Phayung Meesad</i>	450
09:15-09:30 NCCIT2014-195	Forecasting Agricultural Product Prices using Moving Average and Artificial Neural Network Techniques <i>Tanyamai Puchanapakorn and Pudsadee Boonrawd</i>	456
09:30-09:45 NCCIT2014-149	Apply Linear Regression Analysis to Estimate Leaf Area <i>Sattarpoom Thaiparnit, Baramee Osateerakul and Mahasak Ketcham</i>	462
09:45-10:00 NCCIT2014-66	Support Vector Machines For Derivatives Price Prediction <i>Suwat Pinyopan and Boonserm Kijirikul</i>	466
10:00-10:15 NCCIT2014-24	Stock Forecasting with Candlestick Chart and Hidden Markov Models <i>Supakit Nootyaskool, Monsinee Attasiripanya and Wisinee Prasiditichaikul</i>	472
10:15-10:45	Coffee Break	
Data mining Application		
10:45-11:00 NCCIT2014-227	Spatial-Temporal Analysis and Visualization of Twitter Data on Google Earth <i>Isaraporn Vithyaviranont and Veera Muangsin</i>	478
11:00-11:15 NCCIT2014-152	Complex Networks and Online VDO Classification <i>Sureewan Jangjit, Narathip Thongpan, Siriruang Patchuay and Nawaporn Wisitpongphan</i>	484
11:15-11:30 NCCIT2014-159	Hand Written Thai Character Recognition for Mobile Phone <i>Theerapong Boriboon and Patchai Phuangsuan</i>	491
11:30-11:45 NCCIT2014-32	A Feature – Oriented Traceability for Software Product Line Evolution using Vector Space Model <i>Krittika Tipyodsree and Pornsiri Muenchaisri</i>	497
11:45-12:00 NCCIT2014-142	Storing Binary Data to Database using Base128 encoding <i>Chanyaphon Prasomsat and Chaiyapon Khemapatapan</i>	504
12.00-13.00	Lunch	

Friday May 9 th , 2014		
NCCIT2014 Room I		
Feature Selection Techniques, Feature Recognition, Feature Detection, and Gesture Recognition		
Time	Title/Author	Page
13:00-13:15 NCCIT2014-116	Multi-layers Feature Selection Method Using TLiSVM for Dimensionality Reduction <i>Wipawan Buathong and Phayung Meesad</i>	509
13:15-13:30 NCCIT2014-108	Feature Selection and Efficiency Comparison of Classification Techniques for Quail Gender Forecasting from the External Factors of Quail Eggs <i>Jakkarin Suksawatchon, Ureerat Suksawatchon and Pongpat Sing Sri</i>	515
13:30-13:45 NCCIT2014-133	Weighted Features for Online Shopping Fraud Detection <i>Tanasanee Phienthrakul and Nitipat Petmongkonjaras</i>	522
13.45-14.00 NCCIT2014-241	Data Dimension Reduction by Association and Application for Analysis Smart Phone Usage Data <i>Prudtipong Pengsiri, Suriya Pinitkan, Nutchuda Mongkolchart, Nawaporn Wisitpongphan and Phayung Meesad</i>	528
14.00-14.15 NCCIT2014-190	Trend Analysis Model for Information and Communication Technology Regarding the Thai Population Using Data Mining with Factor Loading <i>Natwalai Komkham and Pudsadee Boonrawd</i>	535
14.15-14.30 NCCIT2014-52	Selection Factors Affecting Learning Achievement Following Grad of Students Cluster by Subspace Clustering Algorithms <i>Werayut Pimpaporn and Phayung Meesad</i>	541
14.30-14.45 NCCIT2014-53	Classification of Thai Expressive Speech for Requesting and Ordering using Neural Networks <i>Kanda Rongsawad and Watchara Chatwiriya</i>	547
14.45-15.00 NCCIT2014-94	A Development of Features Detection using Scale Invariant Feature Transform Approach for Herbal Leaves <i>Rakwarinn Wannasin and Krittachai Boonsivanon</i>	553
15.00-15.15 NCCIT2014-158	Time Series Forecasting of Municipal Solid Waste Generation in Bangkok using Artificial Neural Network <i>Paiboon Chatthong, Maleerat Sodanil and Kiattisak Yochanang</i>	559
15.15-15.30 NCCIT2014-136	Development of White Horse Game to Basic English Learning on Android <i>Walongkorn Pasakanon</i>	565
15.30-16.00	Coffee Break	

Friday May 9th, 2014		
NCCIT2014 Room II		
Image and Computer Vision		
Time	Title/Author	Page
09:00-09:15 NCCIT2014-80	Thai Hand Shape Recognition Using HOG - PCA and SVM <i>Anusorn Untao and Sompong Valuvanathon</i>	571
09:15-09:30 NCCIT2014-93	ARtronics: An Object-based Recognition of Augmented Reality for PCB Assembly using SURF Algorithm <i>Krittachai Boonsivanon and Anupap Meesomboon</i>	577
09:30-09:45 NCCIT2014-28	Postural Classification during Television Watching using Kinect <i>Thammarsat Visutarom, Pornchai Mongkolnam and Jonathan Chan</i>	583
09:45-10:00 NCCIT2014-259	An Improved Method for Elementary Cellular Automata Image Encryption Scheme <i>Mongkol Thongkraikaew and Rungrat Wiangsripanawan</i>	589
10:00-10:15 NCCIT2014-223	Feature Extraction For Face Recognition <i>Piyawan Sanpanya, Aumaporn Somthong, Sirapat Chiewhanwattana and Khamron Sunat</i>	595
10:15-10:45	Coffee Break	
10:45-11:00 NCCIT2014-67	The Study of Using CCD Camera to Measure the Lycopene in Tomato by Colors with Artificial Neuron Networks <i>Rapeepong Rattanawaorahirunkul, Sarawut Supannarach and Jitsaran Seekuka</i>	601
11:00-11:15 NCCIT2014-248	Enhancing Dual Bitmap Index with Efficient Encoding <i>Sinate Kimseng and Sirirut Vanichayobon</i>	607
11:15-11:30 NCCIT2014-235	Bitmap Join Index Selection for Data Warehouse under the Space Constraint <i>Hathai Kaewkorn and Sirirut Vanichayobon</i>	613
11:30-11:45 NCCIT2014-191	Applied AR Technique to Present Information of Real Estate House Plans <i>Suntis Jullapho and Nattavee Utakrit</i>	619
11:45-12:00 NCCIT2014-92	Image Watermarking using Joined Wavelet and Time Domain <i>Akarapol Padungdit, Sippawit Sugree and Panwit Tuwanut</i>	625
12.00-13.00	Lunch	

Friday May 9 th , 2014		
NCCIT2014 Room II		
Association Analysis and Decision Support Systems		
Time	Title/Author	Page
13:00-13:15 NCCIT2014-268	Association Rule Discovery from Helpdesk Support System <i>Patharawut Saengsiri, Chanokthamon Suksri and Pornnapa James</i>	631
13:15-13:30 NCCIT2014-151	Application of Genetic Algorithm in Maritime Force Composition: Case Study in Humanitarian Assistance and Disaster Relief Operations <i>Thanapon Boonsanpol and Narongdech Keeratipranon</i>	638
13:30-13:45 NCCIT2014-100	Decision Support System for Selecting Crops and Organic Fertilizers Suitable for Each Other <i>Phanarut Srichetta, Vishit Shuuttha and Thiraphat Pholsri</i>	645
13:45-14:00 NCCIT2014-54	Durable Articles Management System with Rule-Based Expert System: A Case Study on Academic Reserves Center of Khon Kaen University <i>Matupoom Kittisuntaropast, Tanyada Taychatanont, Nipawan Sapet and Nunnapus Benjamas</i>	652
14:00-14:15 NCCIT2014-201	Decision Support Systems for the Export-Import of Thailand <i>Somjate Srinakorn and Maleerat Sodanil</i>	658
14:15-14:30 NCCIT2014-55	Basic Automotive Fault Diagnosis and Repair System <i>Tanchanok Toopkird, Narumol Fengseekom and Nunnapus Benjamas</i>	665
14:30-14:45 NCCIT2014-181	Land Price Forecasting using Data Mining Techniques <i>Kannika Nutchomphu and Maleerat Sodanil</i>	671
14:45-15:00 NCCIT2014-78	Payoff Distribution of ASEAN Power Grid in Optimal Coalition Structure by Shapley Value <i>Benjawan Intara and Panich Sudkhot</i>	677
15:00-15:15 NCCIT2014-50	The Development of Music Information Retrieval System on iOS <i>Kittipun Khantitirat, Chayaphol Prapaipornlert and Kitsuchart Pasupa</i>	683
15:15-15:30 NCCIT2014-5	High School Student Behavior Study Using Association Rule Technique <i>Taweesak Noipasee, Tawa Khampachua and Charun Sanrch</i>	689
15.30-16.00	Coffee Break	

Friday May 9th, 2014		
NCCIT2014 Room III		
Cloud and Network Applications		
Time	Title/Author	Page
09:00-09:15 NCCIT2014-129	A Comparative Study of VoIP Quality between LINE Application and Facebook Using Subjective Assessment <i>Khemika Moosikaphan, Therdpong Daengsi and Pongpisit Wuttidittachotti</i>	694
09:15-09:30 NCCIT2014-134	Task Scheduling Based on Load Balancing in Cloud Computing Using Artificial Bee Colony Algorithm <i>Boonhatai Kruekaew and Warangkhana Kimpan</i>	700
09:30-09:45 NCCIT2014-7	Cost-Based Replacement Policy for Cloud Storage <i>Sumet Siriruangapai and Natawut Nupairoj</i>	706
09:45-10:00 NCCIT2014-59	File Upload Manager via Websocket Protocol for Android Smartphone <i>Worapong Ratanaprasit and Chate Patanothai</i>	713
10:00-10:15 NCCIT2014-258	Performance Evaluation of 3G Networks for Multimedia: A Case Study of Voice Quality Measurement from Skype and Facebook in the Inner City of Bangkok <i>Worawat Akapan, Tuul Triyason, Therdpong Daengsi and Pongpisit Wuttidittachotti</i>	719
10:15-10:45	Coffee Break	
Social Graph		
10:45-11:00 NCCIT2014-150	Analysis of Connecting Computer Network System with Graph Theory A Case Study of Thonburi Commercial College <i>Supaporn Kurdkit, Lawankorn Soimart and Sunantha Sodsee</i>	725
11:00-11:15 NCCIT2014-178	Analysis of Social Networking by Graph Theory: The Case of Faculty Members of Business Administration, Bangkok Thonburi University <i>Chonnikarn Rodmorn and Mathuros Panmuang</i>	731
Ontology		
11:15-11:30 NCCIT2014-154	TCSEdPM: A Domain Ontology for the Teacher Civil Service and Educational Personnel Management <i>Duangporn Jeamumporn, Supot Nitsuwat and Maleerat Sodanil</i>	737
11:30-11:45 NCCIT2014-51	Design Ontologies Technology Research of Teachers and Educational Personnel <i>Titapat Utthakan</i>	743
11:45-12:00 NCCIT2014-40	An Ontology-based Personalized Learning Resource Recommendation for Learning Disability Students <i>Worachat Sukra and Noppamas Pukkhem</i>	749
12.00-13.00	Lunch	

Friday May 9th, 2014		
NCCIT2014 Room III		
Recommender Systems		
Time	Title/Author	Page
13:00-13:15 NCCIT2014-161	The Efficient Rating Prediction Algorithm for Music Recommender System by Using Tendencies Based Algorithm with Micro Profiles <i>Sumet Darapisut and Jakkarin Suksawatchon</i>	755
13:15-13:30 NCCIT2014-257	Consumer Prices Index Prediction using Data Mining Techniques: A Case Study of Sakaeo Province <i>Kittisak Onuean and Phayung Meesad</i>	762
13:30-13:45 NCCIT2014-6	Recommender System for Purchasing Mobile Phones using User-based Collaborative Filtering Approach <i>Phanarut Srichetta and Phonpimon Panthuvapee</i>	768
13.45-14.00 NCCIT2014-183	Recommender Systems for South National Park of Thailand Tourism Using Google Maps API on Smart Phone <i>Chamaiporn Tongkaw and Maleerat Sodanil</i>	774
14.00-14.15 NCCIT2014-27	A Recommendation System to Choose Study Program based on Repertory Grid <i>Dechawut Wanichsan and Sathit Suwannawach</i>	780
Feature Selection Techniques, Feature Recognition, Feature Detection, and Gesture Recognition		
14.15-14.30 NCCIT2014-110	Defect Detection in Textile Fabrics with Edge Detection and Support Vector Machines <i>Thodsaphon Prachpreecha and Sorakis Srikasem</i>	786
14.30-14.45 NCCIT2014-61	A Comparative Study in Feature Extraction Methods of Peg-fixed Hand Geometry Device <i>Jakkree Srinonchat and Apiwat Sawatdirat</i>	792
14.45-15.00 NCCIT2014-135	Real-Time Lane Detection for Driving Using Embedded System <i>Kittisak Thongyuan and Mahasak Ketcham</i>	798
15.00-15.30	<i>Coffee Break</i>	

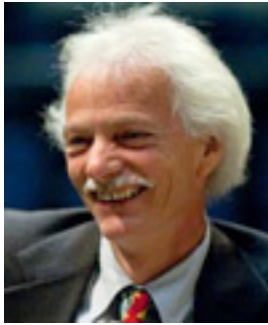
Friday May 9th, 2014		
NCCIT2014 Room IV		
Human-Computer Interaction		
Time	Title/Author	Page
09:00-09:15 NCCIT2014-210	Design Virtual Keyboard Layout Using Evolutionary Computation <i>Veerawan Janthanasub, Phayung Meesad and Metha Sirigool</i>	804
09:15-09:30 NCCIT2014-68	A User Action Events Timestamp Tool on Windows Application <i>Rapheephong Sriwiboon and Chate Patanothai</i>	810
09:30-09:45 NCCIT2014-234	An Assessment of Effectiveness of Adaptive Interfaces on WWW for Visually Impaired Users <i>Panida Tunyanon and Worasit Choochaiwattana</i>	816
09:45-10:00 NCCIT2014-89	A Development of Health-related Websites for Aging Users Based on TWCAG 2010 <i>Duangjai Noolek and Thippaya Chintakovid</i>	822
10:00-10:15 NCCIT2014-31	Implementation of edit distance with Thai language <i>Metawee Jindarak and Supakit Nootyaskool</i>	828
10:15-10:45	Coffee Break	
10:45-11:00 NCCIT2014-222	A Eleven-Syllable Poetry Analysis using Word Segmentation and Rule-Based <i>Isranol Polparee and Pudsadee Boonrawd</i>	834
Software Engineering		
11:00-11:15 NCCIT2014-217	Factors Influencing Intention to use Cloud-Based Project Management Software in Private Organization <i>Supitchaya Janplong, Maneewan Supatimassaro and Kanokwan Atcharyachanvanich</i>	840
11:15-11:30 NCCIT2014-84	A Comparison Study of Test Effort between Manually Revised Test Case and Test Case Refactoring for Pattern-Based Software Development <i>Jaturada Deeying and Assadaporn Sapsomboon</i>	846
11:30-11:45 NCCIT2014-33	Identifying Common Asset Candidates in Source Code by Comparing Software Testing Paths <i>Sithipon Limchaichada and Pornsiri Muenchaisri</i>	853
12.00-13.00	Lunch	

Friday May 9th, 2014		
NCCIT2014 Room IV		
Management Information System		
Time	Title/Author	Page
13:00-13:15 NCCIT2014-204	Management Information System for Sugar Cane Garden <i>Natthawan Phoonson, Pariwat Kumkunkum and Autthapon Hutthapanit</i>	859
13:15-13:30 NCCIT2014-260	The Technique Using a Widget to Control The Change Management of Software: A Case Study of Information Systems for Education Quality Assurance: Praboromarajchanok Institute Executive Information System : PIEiS <i>Athita Onuean</i>	865
13:30-13:45 NCCIT2014-166	Prototype of Decision Support System for Emergency Medical Triage Protocol and Criteria Based Dispatch A Case Study of National Institute for Emergency Medicine <i>Piradee Chalompattanapong and Pongpisit Wuttidittachotti</i>	871
13:45-14:00 NCCIT2014-4	An Analysis of Factors Influencing Hybrid Self-Regulated and Collaborative Learning for End-User Training <i>Sirikorn Krompho, Montean Rattanasiriwongwut and Nalinpat Porrawatpreyakorn</i>	877
14:00-14:15 NCCIT2014-192	On the Application of Solr software for Helpdesk and Retrieval issue TOT iptv <i>Krit Sirisup and Nattavee Utakrit</i>	884
Business Intelligence		
14:15-14:30 NCCIT2014-194	NAV Forecasting Model Using Data Mining and Artificial Neural Network <i>Thararat Udomchaibanjerd and Pudsadee Boonrawd</i>	890
14:30-14:45 NCCIT2014-225	Forecast the Relationship of the Customer to Buy the Long Term Fund (LTF) <i>Chaowat Prattanasangkusol and Maleerat Sodanil</i>	896
14:45-15:00 NCCIT2014-189	Cooperative Financial Analysis with CAMELS Analysis and Business Intelligence <i>Patsapol Prommat and Pudsadee Boonrawd</i>	902
15:00-15:15 NCCIT2014-60	An Membership System for Android Smartphone with NFC Card Emulation Mode <i>Nattaphon Saelim and Chate Patanothai</i>	908
15.30-16.00	Coffee Break	

Friday May 9th, 2014		
NCCIT2014 Room V		
Computer Education		
Time	Title/Author	Page
09:00-09:15 NCCIT2014-71	The Learning Achievement of the Students Through the Project-based Learning on the Cooperate with the Community on Information Management Course <i>Apichat Lagdee</i>	914
09:15-09:30 NCCIT2014-35	A Development of Computer-Assisted Instruction Program on the Topic of “Estimation of Vertical Cup-to-Disc Ratio from Retinal Fundus Images” for Eye Residents <i>Narumol Nuanchawee and Nongluk Covavisaruch</i>	920
09:30-09:45 NCCIT2014-115	Factors Affecting Autistic Childern in Parents’ Perspectives and the Use of Multimedia Art Game Set to Create Color Learning Process Divided by Groups of Symptoms <i>Kanisorn Jeekratok</i>	927
09:45-10:00 NCCIT2014-65	Creating a Math Class Grade 1 with CFG: Application on Tablet PCs with Automatic Analysis of Learning Outcomes <i>Wachira Khaengngan and Sunisa Rimcharoen</i>	934
10:00-10:15 NCCIT2014-15	The Development of Learning English Vocabulary with Little Red Riding Hood on Tablet <i>Nattavadee Hongboonmee and Phichitra Siritwat</i>	940
10:15-10:45	Coffee Break	
10:45-11:00 NCCIT2014-10	The Applications of eDLTV Media on Science Learning Strand through the Application Development for Learning on the Tablet <i>Anukul Butprom and Worapapha Arreerard</i>	946
11:00-11:15 NCCIT2014-126	A Development of 3D Game Computer Assisted Instruction Computer for Grade 2 Students. Case Study at Watladplakao School <i>Charinthorn Aumgril</i>	952
11:15-11:30 NCCIT2014-173	Development of Online Participatory Teaching Media Access via 2D Barcode For Measure of Intention in Learning Context and Enhancing Practice Package Software Skill <i>Ubonrat Sirisukpoca and Paisan Simalaotao</i>	957
11:30-11:45 NCCIT2014-95	The Behavior Analysis on the Applying Major Selection and the Comparison of Model to Forecast the Number of New Students Using Data Mining Technique <i>Teerapong Sungsi</i>	963
11:45-12:00 NCCIT2014-163	The Visualization of Student Information in the Schools of Nakhon Ratchasima Provincial Administrative Organization <i>Nikhom Wuttiya, Khamron Sunat and Sirapat Chiewchanwattana</i>	969
12.00-13.00	Lunch	

Friday May 9 th , 2014		
NCCIT2014 Room V		
Management Information System		
Time	Title/Author	Page
13:00-13:15 NCCIT2014-261	A Framework of Team Recruitments with Feature Selection and Competency-based Management <i>Julaluk Watthananon</i>	975
13:15-13:30 NCCIT2014-226	The Development of Business Intelligence to Production Planning in The Organization <i>Prapada Pokumapichai and Montean Rattanasiriwongwut</i>	982
13:30-13:45 NCCIT2014-125	Development of a Graduate Level Student Relationship Management Model <i>Jirapong Pongpilasarn and Sakchai Tangwannawit</i>	988
13.45-14.00 NCCIT2014-207	Affecting Analysis Factor to the IT Competence of PTT PLC., Officers <i>Nitikorn Intarasod and Montean Rattanasiriwongwut</i>	994
14.00-14.15 NCCIT2014-219	The Guidance for security operation of information technology using ISO/IEC 27001 Standard A case study of Electronic Government Agency: EGA (Public Organization) <i>Ratchaporn Suphap and Sakchai Tangwannawit</i>	1000
14.15-14.30 NCCIT2014-143	Application of ITIL Standard for Internet Service Support System <i>Benjaporn Paripinidchai and Mahasak Ketcham</i>	1006
14.30-14.45 NCCIT2014-8	Decision Support System for Metropolitan Waterworks Authority Manager <i>Aiyapan Eagobon and Nuengwong Tuaycharoen</i>	1012
14.45-15.00 NCCIT2014-171	The Zachman Framework for Graduate Thesis and Dissertation Management System in Physics at Mahidol University <i>Sunti Preechawit and Phumin Kirawanich</i>	1018
15:00-15:15 NCCIT2014-164	Mushrooms Cube Tracking and Traceability System <i>Anuwat Jaidee and Pusadee Seresangtakul</i>	1024
15.15-15.30 NCCIT2014-228	Employee Retention Model base on Employee Engagement <i>Teeraporn Kamwong and Sakchai Tangwannawit</i>	1030
15.30-16.00	<i>Coffee Break</i>	

KEYNOTE SPEAKERS



Prof. Dr. Peter Kropf

Dean of Faculty of Sciences

University of Neuchâtel, Switzerland

About the Speaker:

Prof. Dr. Peter Kropf received his M.Sc. (Mathematics) and his Ph.D. degrees (Computer Science) from the University of Bern, Switzerland. From 1994 to 1999, he was an assistant and associate professor at Laval University, Quebec, Canada. From 2000 to 2003, he was appointed as an associate professor at the department of computer science and operations research (DIRO) at University of Montreal, Canada. Since October 2003, he is a Professor and head of Department of Computer Science at the University of Neuchâtel, Switzerland. Peter Kropf is currently the dean of the Faculty of Sciences. He is also associated researcher at the Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation, Montreal. He published about a hundred research papers in the fields of distributed and parallel systems, simulation and optimization.

Topic: Wireless Mesh Networks and Cloud Computing for Real Time Environmental Simulations

Abstract

Predicting the influence of drinking water pumping on stream and groundwater levels is essential for sustainable water management. Given the highly dynamic nature of such systems any quantitative analysis must be based on robust and reliable modeling and simulation approaches. The paper presents a wireless mesh-network framework for environmental real time monitoring integrated with a cloud computing environment to execute the hydrogeological simulation model. The simulation results can then be used to sustainably control the pumping stations. The use case of the Emmental catchment and pumping location illustrates the feasibility and effectiveness of our approach even in harsh environmental conditions.

**Assoc. Prof. Dr. Hiroya Takamura**

Advanced Information Processing Division

Precision and Intelligence Laboratory

Tokyo Institute of Technology

About the Speaker:

Hiroya Takamura received his B.E. and M.E. from the University of Tokyo in 1997 and 2000, respectively. He received his Dr.Eng. from the Nara Institute of Science and Technology in 2003. He was an assistant professor at the Tokyo Institute of Technology from 2003 to 2010. He is currently an associate professor at Tokyo Institute of Technology. His current research interests include sentiment analysis and text summarization.

Topic: Mathematical Models for Text Summarization**Abstract**

Text summarization is a task of automatically generating a short and concise document that represents the core content of the given documents. It has long been a major research task in the natural language processing field and is becoming more and more important with the recent rapid growth of amount of text data.

In this keynote speech, I am going to talk about the recent advances of text summarization methods. After introducing basics knowledge pertaining to summarization, I will present some well-known summarization methods, starting with some conventional methods, and going on to more recent methods such as graph-based methods, mathematical models based on combinatorial optimization problems. In particular, I will elaborate on the maximum coverage summarization model [1], in which the generated summary is supposed to cover as many word types in the given documents as possible. I will also talk about how machine-learning techniques from simple classification methods through the structured output learning framework, are applied to text summarization.

I will also give a view on each piece of the major work in this area, from the perspective of what part of the summarization techniques each of the major efforts contributes to, i.e., model, parameter estimation, and decoding. Following this view, I will discuss what is missing in this area and give future directions in text summarization research.

การจำแนกต้นไม้ตัดสินใจสำหรับชุดข้อมูลไม่สมดุลโดยใช้น้ำหนักต่างกันบนข้อมูลสังเคราะห์

Decision Tree Classification of Imbalanced Data Sets Using Different Weight On Synthesized Data

สุรพงษ์ เชี่ยวสกุลวัฒนา (Suraphong Cheawsakunwattana)¹ และสุกรี สินธุภิญโญ (Sukree Sinthupinyo)²

^{1,2}ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

suraphong_ch@hotmail.com, sukree@cp.eng.chula.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อสร้างต้นไม้ตัดสินใจสำหรับชุดข้อมูลไม่สมดุล โดยใช้วิธี C4.5 เป็นพื้นฐาน และได้ปรับการหาค่าเอนโทรปีใหม่ เพื่อให้จำแนกข้อมูลกลุ่มน้อยได้ดีขึ้น สำหรับการจำแนกข้อมูลแบบสองกลุ่ม เลือกชุดข้อมูลไม่สมดุลจำนวน 8 ชุดข้อมูล มาทำการทดลอง ทำการทดสอบแบบไขว้ข้ามสิบกลุ่ม โดยกำหนดนัยสำคัญทางสถิติของค่าความแม่นยำทั้งหมดที่ระดับ 0.1 และ 0.05 ตามลำดับ การทดลองพบว่าวิธีการที่นำเสนอสามารถจำแนกตัวอย่างกลุ่มน้อยได้ดี เมื่อใช้น้ำหนักต่างกันบนข้อมูลสังเคราะห์ โดยวิธีการที่นำเสนอ (AMOTE) ให้ค่าความระลึก (Recall) ในข้อมูลกลุ่มน้อยได้ดีกว่าวิธี C4.5-SMOTE จำนวน 6 ชุดข้อมูล และประสิทธิภาพโดยรวมให้ค่าที่ดีกว่าวิธี C4.5 และ SMOTE จำนวน 7 ชุดข้อมูล จากข้อมูลที่น่ามาทดสอบทั้งหมด 8 ชุดข้อมูล

คำสำคัญ: ชุดข้อมูลไม่สมดุล ข้อมูลกลุ่มมาก ข้อมูลกลุ่มน้อย การสุ่มเพิ่มข้อมูลกลุ่มน้อย ข้อมูลสังเคราะห์

Abstract

This study is aimed at constructing a decision tree for an imbalanced data set using C4.5 basis with a new entropy adjustment in order to enhance the minority classification. For the two classes of data classification, 8 imbalanced data sets were selected and tested using the 10 - fold Cross Validation which determined the statistical significance of the total precision at 0.1 and 0.05. The results show that the presented method were able to identify minority class well. When using the different weight on synthetic data, the result shows that AMOTE, our method, obtained higher Recall in the

minority data was higher than the C4.5 SMOTE for 6 data sets and the overall performance was better than that of the C4.5 SMOTE for 7 data sets from 8 datasets tested in this experiment.

Keyword: Imbalanced data sets, Majority Class, Minority Class, SMOTE, Synthesized data

1. บทนำ

ปัญหาข้อมูลไม่สมดุล (Class Imbalance) เป็นปัญหาที่เกิดขึ้นกับข้อมูลหลายๆ ประเภท ข้อมูลที่เกิดขึ้นโดยทั่วไปจะแบ่งออกเป็น 2 ประเภท คือ ข้อมูลกลุ่มน้อย (Minority class) ซึ่งเป็นข้อมูลที่มีอยู่จำนวนน้อยในชุดข้อมูล ส่วนอีกประเภทหนึ่งคือ ข้อมูลกลุ่มมาก (Majority class) จะเป็นกลุ่มข้อมูลที่มีอยู่จำนวนมากในชุดข้อมูล [1]

จากสาเหตุของกรณีที่มีข้อมูลที่ไม่สมดุลเกิดขึ้น เมื่อต้องการทำการจำแนกประเภทข้อมูล ผลลัพธ์ที่ได้เมื่อทำการจำแนกข้อมูลจะถูกจำแนกไปในกลุ่มข้อมูลกลุ่มมาก

วิธีการจำแนกประเภทข้อมูล จากงานวิจัยที่ผ่านมาหลายงานวิจัยได้ให้ความสนใจในการจำแนกข้อมูลที่ไม่สมดุล และตัวอย่างกลุ่มน้อย มีการออกแบบเอนโทรปีใหม่สำหรับต้นไม้ตัดสินใจ เพิ่มตัวปรับยัดเกาะ (Adhesively Modified Entropy: AMIE) [2] มีวัตถุประสงค์เพื่อจำแนกกลุ่มตัวอย่างข้างน้อยให้ดีขึ้น นอกจากนี้ได้มีการพัฒนาจากเทคนิคหรือวิธีการสุ่มเพิ่มตัวอย่างกลุ่มน้อย (SMOTE) [3] และได้้นำเทคนิคนี้มาพัฒนาต่อ โดยทำการสุ่มเพิ่มข้อมูลกลุ่มน้อยและลดข้อมูลกลุ่มมาก (Synthetic Minority Over-sampling and Under-sampling Techniques for class imbalanced problem: SMOUTE) [4] เป็น

ขั้นตอนเตรียมข้อมูลก่อนทำการนำมาจำแนกประเภทด้วยวิธีต่างๆ

งานวิจัยฉบับนี้จึงนำเสนอวิธีการจำแนกข้อมูลไม่สมดุล โดยใช้น้ำหนักที่ต่างกันบนข้อมูลที่ถูกละทิ้งขึ้น เพื่อทำการจำแนกข้อมูลที่ไม่สมดุล และให้ความสนใจในข้อมูลกลุ่มน้อย และเพื่อเพิ่มประสิทธิภาพในการจำแนกข้อมูลกลุ่มน้อยได้ดีขึ้น

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ต้นไม้ตัดสินใจ (Decision Tree)

ต้นไม้ตัดสินใจ (Decision Tree) คือ แบบจำลองที่มีลักษณะเหมือนโครงสร้างของต้นไม้ที่มีลักษณะเป็นต้นไม้กลับหัว แต่ละโหนดจะแสดงคุณลักษณะ (attribute) ส่วนกึ่งจะแสดงผลในการทดสอบ และโหนดใบแสดงคลาส ซึ่งเป็นผลลัพธ์ที่ได้จากการจำแนกข้อมูล [5]

2.2 อัลกอริทึม ID3

เป็นอัลกอริทึมที่ใช้หลักการของการสร้างต้นไม้ตัดสินใจ ซึ่งนิยมนำมาใช้ในการจำแนกข้อมูล ซึ่งจะมีการใช้ค่าเกณฑ์มาตรฐาน (Information Gain) เป็นตัวคัดเลือก โดยหาค่าเกณฑ์มาตรฐานทุกๆ คุณลักษณะในข้อมูล แล้วเลือกค่าเกณฑ์มาตรฐานที่มีค่ามากที่สุดมาเป็นโหนดเริ่มต้น โดยค่าเกณฑ์มาตรฐานจะคำนวณจากการหาค่าเอนโทรปีก่อนแบ่งแยกกับด้วยค่าเอนโทรปีหลังแบ่งด้วยคุณลักษณะที่พิจารณา ซึ่งสมการต่างๆ คือ

$$\text{Entropy}(D) = - \sum_{i=1}^n P_i \log P_i \quad (1)$$

$$\text{Entropy}_A(D) = - \sum_{i=1}^n \frac{|D_j|}{|D|} \times \text{Entropy}(D_j) \quad (2)$$

Entropy(D) คือ ค่า Entropy ของเซต (D) ข้อมูลทั้งหมด
 P_i คือ ความน่าจะเป็นที่ record ในเซต D เป็นคลาสของ C_i
 $\text{Gain}(A) = \text{Entropy}(D) - \text{Entropy}_A(D)$ (3)

จากสมการข้างต้นจะทำให้สามารถหาค่าของข้อมูลที่เป็นกลุ่มเดียวกันได้ ข้อมูลชุดใดมีค่า Gain มากที่สุด ชุดข้อมูลนั้นจะถูกเลือกเป็นโหนดราก แล้วทำเช่นนี้กับข้อมูลที่เหลือจนได้ต้นไม้ที่สมบูรณ์

2.3 อัลกอริทึม C 4.5

เป็นอัลกอริทึมที่พัฒนาต่อมาจาก ID3 โดย Quinlan [7] ซึ่งเป็นการแก้ปัญหากรณีที่มีข้อมูลในชุดข้อมูลนั้นยังมีการกระจายตัวอยู่มาก ยังไม่เป็นกลุ่มเดียวกัน ทำให้การจำแนกเกิด

ความเอนเอียง จึงเพิ่มค่าสารสนเทศการแบ่งแยก (Split Information) คำนวณได้จากสมการ

$$\text{SplitInfo}_A(D) = - \sum_{i=1}^n \frac{|D_j|}{|D|} \times \log_2 \left(\frac{|D_j|}{|D|} \right) \quad (4)$$

จากการที่ค่าสารสนเทศการแบ่งแยก (SplitInformation) จะบอกถึงลักษณะการกระจายของข้อมูล วิธีแก้ปัญหาค่าความเอนเอียง จึงนำค่าสารสนเทศในการแบ่งแยกหารด้วยค่ามาตรฐานเกณฑ์ จะทำให้ได้ค่ามาตรฐานอัตราส่วนเกณฑ์ คำนวณได้จากสมการที่ (4)

$$\text{GainRatio}(A) = \frac{\text{Entropy}(D) - \text{Entropy}_A(D)}{\text{SplitInfo}_A(D)} \quad (5)$$

เมื่อคำนวณค่ามาตรฐานอัตราส่วนเกณฑ์ (GainRatio) ได้แล้วเลือกค่าที่มีค่าสูงที่สุดมาเป็นโหนดเริ่มต้น และทำการสร้างโหนดในระดับถัดลงมา โดยใช้คุณลักษณะ (attribute) ที่เหลือ จนได้ต้นไม้ที่สมบูรณ์ หรือข้อมูลมีการกระจายน้อยที่สุดหรือเป็นประเภทเดียวกัน

2.4 การสุ่มเพิ่มตัวอย่างกลุ่มน้อย (Synthetic Minority Over-sampling Technique – SMOTE)

เป็นเทคนิคหรือวิธีการหนึ่งที่ช่วยแก้ปัญหา กรณีที่ต้องการจำแนกข้อมูลไม่สมดุล ซึ่งเป็นข้อมูลที่มีจำนวนตัวอย่างแตกต่างกันมากในแต่ละคลาส เมื่อเรานำข้อมูลที่มีจำนวนแตกต่างกันระหว่างคลาส มาทำการจำแนกประเภท ผลลัพธ์ที่ได้จะทำให้มีการเรียนรู้แต่ข้อมูลกลุ่มมาก และเมื่อทำการจำแนกประเภท ก็จะจำแนกไปในข้อมูลกลุ่มมาก ด้วยเหตุนี้การแก้ปัญหาชุดข้อมูลที่สมดุล จึงเลือกใช้วิธีการเพิ่มจำนวนข้อมูลกลุ่มน้อยขึ้น โดยการสังเคราะห์ข้อมูลขึ้น เฉพาะในข้อมูลกลุ่มน้อย เพื่อให้การจำแนกประเภทในข้อมูลกลุ่มน้อยดีขึ้น เป็นเทคนิคหนึ่งที่ใช้เพื่อเตรียมข้อมูลก่อนที่จะนำไปทำการจำแนกข้อมูลด้วยวิธีการ C4.5 [8][9][10][11]

2.5 งานวิจัยที่เกี่ยวข้อง

อุไรรัตน์ [2] ได้นำเสนอวิธีการที่จะทำการจำแนกกลุ่มน้อยอย่างกลุ่มน้อยให้ดีขึ้น ด้วยการใช่วิธี C4.5 และได้ออกแบบเอนโทรปีใหม่สำหรับต้นไม้ตัดสินใจ เรียกว่า เอนโทรปีตัวปรับแบบยึดเกาะ (Adhesively Modified Entropy: AMIE) ทำการเปรียบเทียบกับอัลกอริทึม C4.5 เอนโทรปีแบบอสมมาตร (AE) และเอนโทรปีแบบออกจากศูนย์กลาง (OCE) ซึ่งวิธีการ AMIE ให้ผลในการจำแนกกลุ่มตัวอย่างกลุ่มน้อยได้ดี แต่ข้อจำกัด

สำหรับงานวิจัยฉบับนี้ในการจำแนกข้อมูลที่ไม่สมดุล คือ ถ้าจำนวนคุณลักษณะ (Attribute) ของข้อมูลน้อยเกินไป จะทำให้ประสิทธิภาพในการทำงานลดลง และจะทำให้ค้นไม่ตัดสินใจที่ได้มีขนาดใหญ่ และการแทนความรู้ในรูปแบบของกฎจากค้นไม่ตัดสินใจจะมีจำนวนมากด้วย

ปณต [4] เสนอเทคนิคการสุ่มเพิ่มตัวอย่างข้างน้อยสังเคราะห์และเทคนิคการสุ่มลดตัวอย่างข้างมากสำหรับปัญหาความไม่สมดุลระหว่างกลุ่ม (SMOUTE) ซึ่งเป็นกระบวนการจัดเตรียมข้อมูลก่อนที่จะทำการจำแนกข้อมูล โดยจะทำการเพิ่มจำนวนข้อมูลกลุ่มน้อยด้วยเทคนิค (SMOTE) และลดจำนวนข้อมูลกลุ่มมาก จนทั้ง 2 กลุ่มมีขนาดใกล้เคียงกัน โดยใช้ตัวจำแนก 3 วิธี คือ C4.5 การแบ่งประเภทแบบเบย์ และเพอร์เซ็ปตรอนหลายชั้น ซึ่งในการทดลองวิธี SMOUTE ให้ผลในการทำนายข้อมูลได้ดีกว่าวิธี SMOTE ในข้อมูลที่มีขนาดใหญ่ แต่ข้อจำกัดสำหรับงานวิจัยนี้ คือถ้าใช้วิธีการ SMOUTE จะเหมาะกับตัวจำแนกประเภทที่ใช้หลักการการแบ่งข้อมูลแบบเชิงเส้น (Linear) และข้อมูลแบบไม่เชิงเส้น Non-Linear

Boonchuay [11] ได้เสนออัลกอริทึมที่ใช้ค่าที่แตกต่างกันของกลุ่มน้อยที่จะทำให้มาตรฐานอัตราส่วนเกินสูงขึ้น จะเลือกค่าเอนโทรปีของข้อมูลกลุ่มน้อยที่ค่าที่ต่ำ โดยแก้ปัญหาชุดข้อมูลที่ไม่สมดุล ซึ่งอาจเกิดการจำแนกตัวอย่างที่ผิดได้ในกรณีที่ตัวอย่างกลุ่มน้อยมีอยู่จำนวนน้อยมาก โดยให้แนวคิดว่าการผิดของการไปทำการลดตัวอย่างกลุ่มมากที่มีอยู่อาจจะไปลดความสำคัญของตัวอย่างนั้นลงไปด้วย

3. วิธีการดำเนินการวิจัย

ในการศึกษาวิจัยครั้งนี้ เสนอวิธีการในการจำแนกค้นไม่ตัดสินใจด้วยวิธีการ C4.5 ซึ่งใช้น้ำหนักต่างกันบนข้อมูลสังเคราะห์ในชุดข้อมูลที่ไม่สมดุล ได้เลือกชุดข้อมูลที่ไม่สมดุลจาก UCI Machine Learning Repository [12] จำนวน 8 ชุดข้อมูล แล้วดำเนินการตามขั้นตอนดังต่อไปนี้

3.1 การเตรียมข้อมูล

เป็นกระบวนการเตรียมข้อมูลก่อนที่จะนำข้อมูลเข้าสู่กระบวนการจำแนกข้อมูล [13] ด้วยการแทนข้อมูลไม่ทราบค่าและการแปลงข้อมูลชนิดตัวเลขให้เป็นข้อมูลชนิดไม่ต่อเนื่อง แล้วจัดเก็บในรูปแบบ arff file สำหรับใช้ในการทดสอบ ซึ่งรายละเอียดของข้อมูลทั้ง 8 ชุด แสดงในตารางที่ 1

ตารางที่ 1: รายละเอียดข้อมูลไม่สมดุลที่ใช้ในการวิจัยซึ่งเป็นข้อมูลที่นำมาจาก UCI Machine Learning Repository และเมื่อใช้เทคนิค SMOTE เพิ่มข้อมูลกลุ่มน้อย

ชื่อชุดข้อมูล	ชุดข้อมูล	คุณลักษณะ	จำนวนกลุ่มมาก	จำนวนกลุ่มน้อย	จำนวนทั้งหมด
Post-operative	Original	8	66	24	90
	SMOTE	8	66	48	114
Liver	Original	6	200	145	345
	SMOTE	6	200	290	490
Ionosphere	Original	34	225	126	351
	SMOTE	34	255	252	477
glasses	Original	9	197	17	214
	SMOTE	9	197	34	231
ecoli	Original	7	301	35	336
	SMOTE	7	301	70	371
Flags	Original	28	177	17	194
	SMOTE	28	177	34	211
Hepatitis	Original	19	123	32	155
	SMOTE	19	123	64	187
Credit-g	Original	20	700	300	1000
	SMOTE	20	700	600	1300

3.2 วิเคราะห์และออกแบบการทดลอง

ทำการออกแบบและปรับสมการเอนโทรปีของอัลกอริทึม C4.5 เพื่อให้สามารถใส่ค่าน้ำหนักที่ต่างกันบนข้อมูลกลุ่มน้อยที่มีอยู่เดิมและข้อมูลกลุ่มน้อยที่ถูกสังเคราะห์ขึ้นมาใหม่ เรียกว่าวิธีการ AMOTE ซึ่งสมการที่ได้ออกแบบมีดังนี้

$$\begin{aligned} \delta &= \left(\frac{S_1}{S_1+S_2+S_3} \right) \\ \epsilon &= \left(\frac{S_2}{S_1+S_2+S_3} \right) \\ \gamma &= \left(\frac{S_3}{S_1+S_2+S_3} \right) \\ \varphi &= \left(\frac{S_2+S_3}{S_1+S_2+S_3} \right) \end{aligned}$$

โดย S_1 คือ จำนวนตัวอย่างในกลุ่มมาก

S_2 คือ จำนวนตัวอย่างในกลุ่มน้อยที่มีอยู่เดิม

S_3 คือ จำนวนตัวอย่างในกลุ่มน้อยที่เพิ่มใหม่จากการ SMOTE

β คือ ค่าน้ำหนักที่ให้กับตัวอย่างกลุ่มน้อย

α คือ จำนวนเท่า ในการให้ค่าน้ำหนักกับตัวอย่างกลุ่มน้อย

AMOTE คือ เอนโทรปีที่ออกแบบใหม่ โดยให้ค่าน้ำหนักต่างกันบนข้อมูลกลุ่มน้อยที่ถูกสังเคราะห์ขึ้นมาใหม่ (β) ด้วยเทคนิค SMOTE เมื่อกำหนดให้ $\alpha=2$ จะทำให้ผลในการจำแนกตัวอย่างได้ผลออกมาเท่ากับ วิธีการ C4.5 แบบพื้นฐานที่ไม่ได้

ปรับค่า เมื่อใช้ข้อมูลที่น่าไปทำการสังเคราะห์ข้อมูลกลุ่มน้อย ด้วยเทคนิค SMOTE ที่ทำการสังเคราะห์ข้อมูลกลุ่มน้อยขึ้นใหม่ 100%, k=5 โดยสมการของ AMOTE มีดังนี้

Entropy ก่อนแบ่งแยก (AMOTE)

$$\text{Entropy (AMOTE)} = - \delta \log_2 \delta - [(\alpha\beta) \times (\epsilon \log_2 \epsilon) - [(\beta) \times (\gamma \log_2 \gamma)]] \quad (6)$$

โดย β หาค่าได้จาก $\beta = \log_{[(\epsilon^{\alpha\epsilon}) \times (\gamma^{\gamma})]} \varphi^{\varphi}$

Entropy หลังแบ่งแยกด้วยคุณลักษณะที่พิจารณา (AMOTE)

$$\text{Entropy}_A (D_j \text{ AMOTE}) = - \sum_{i=1}^n \frac{|D_{ji}|}{|D_j|} \times \text{Entropy}(D_j \text{ AMOTE}) \quad (7)$$

3.3 วิธีทดลอง

หลังจากออกแบบเอนโทรปีที่ใช้ในการทดลอง จากการปรับสมการเอนโทรปีของอัลกอริทึม C4.5 (AMOTE) โดยจะเปรียบเทียบกับอัลกอริทึม C4.5 ที่ใช้ข้อมูลที่ทำกร SMOTE ขึ้น ดังสมการที่ 6 ส่วนในการทดสอบแสดงดังตาราง

ตารางที่ 2: การทดสอบเปรียบเทียบระหว่าง 2 เทคนิค

เทคนิค	ชุดข้อมูลสอน	ชุดข้อมูลทดสอบ
C4.5 (SMOTE)	SMOTE	SMOTE
C4.5 (AMOTE)	SMOTE	SMOTE

SMOTE ในชุดข้อมูลสอนและข้อมูลทดสอบ คือ ชุดข้อมูลที่น่ามาจาก UCI Machine Learning Repository [12] และถูกสังเคราะห์ข้อมูลกลุ่มน้อยขึ้นด้วยเทคนิค SMOTE โดยกำหนดให้ ข้อมูลกลุ่มน้อยถูกสังเคราะห์เพิ่มขึ้น 100% จากชุดข้อมูลต้นฉบับที่ได้ทำการแปลงค่าต่างๆ แล้ว และ k = 5 [14]

ในการทดลองจะทำการกำหนดน้ำหนักให้กับข้อมูลกลุ่มน้อยที่ต่างกัน จากการหาค่า β เพื่อหาความความเที่ยงในการทำนายกลุ่มน้อยได้ดีที่สุดเมื่อเปรียบเทียบกับวิธี C4.5 แบบพื้นฐานที่ถูกสังเคราะห์ข้อมูลกลุ่มน้อย ด้วยเทคนิค SMOTE

ทำการทดสอบแบบไขว้ข้ามสิบกลุ่ม เพื่อวิเคราะห์ความถูกต้องในการจำแนกแบบจำลองที่ออกแบบไว้

เนื่องจากงานวิจัยฉบับนี้สนใจในชุดข้อมูลไม่สมดุล เพื่อต้องการเพิ่มประสิทธิภาพในการจำแนกกลุ่มน้อยที่ดีขึ้น จึงได้ทำการทดลองเปรียบเทียบเฉพาะค่าความระลึก ค่าความเที่ยง และค่าเอฟ ในกลุ่มข้อมูลที่มีจำนวนน้อยด้วยค่าเฉลี่ยการทดสอบไขว้ข้ามทั้ง 10 ครั้ง ในแต่ละชุดข้อมูล

3.4 ผลการดำเนินงาน และอภิปรายผล

จากการทดลองกับชุดข้อมูลทั้งหมด 8 ชุดข้อมูล ได้ผลค่าความแม่นยำ (Accuracy) ในการทำนายแต่ละกลุ่ม ดังตาราง

ตารางที่ 3: ผลการจำแนกประเภทข้อมูลด้วยวิธี C4.5 ที่ใช้ข้อมูลที่ได้จากการ SMOTE มาทดสอบ กับวิธี C4.5 - AMOTE ที่ได้ออกแบบเอนโทรปีแบบใหม่ และใช้ข้อมูลจากการ SMOTE มาทดสอบด้วยวิธี ไขว้ข้ามสิบกลุ่ม

ชุดข้อมูล	ความแม่นยำทั้งหมด		ความแม่นยำกลุ่มมาก		ความแม่นยำกลุ่มน้อย	
	SMOTE	AMOTE	SMOTE	AMOTE	SMOTE	AMOTE
post-operative ($\alpha = 8$)	68.80	69.39	50.50	50.50	81.90	82.86
Liver ($\alpha = 2.8$)	70.41	73.47*	49.50	56.50*	84.83	85.17
Ionosphere ($\alpha = 5.8$)	88.49	90.97*	86.19	90.22**	90.54	91.65
Glasses ($\alpha = 5.7$)	83.86	85.56**	30.83	33.33	92.82	94.32**
ecoli ($\alpha = 6.8$)	90.03	90.56*	78.57	78.57	92.70	93.35*
Flags ($\alpha = 3.9$)	88.10	90.51*	64.17	70.00	92.61	94.35*
Hepatitis ($\alpha = 2.05$)	80.70	81.81	75.00	75.00	83.72	85.38
Credit-g ($\alpha = 4$)	71.08	71.77	68.67	67.83	73.14	75.14**

* หมายถึงนัยสำคัญทางสถิติของค่าความแม่นยำทั้งหมดที่ระดับ 0.1

** หมายถึงนัยสำคัญทางสถิติของค่าความแม่นยำทั้งหมดที่ระดับ 0.05

ผลจากการจำแนกประเภทข้อมูลด้วยวิธี C4.5 เมื่อใช้ชุดข้อมูลที่ทำกร SMOTE กับวิธี C4.5 ด้วยเทคนิค AMOTE ความแม่นยำวิธี C4.5 – AMOTE มากกว่า วิธี C4.5 – SMOTE จำนวน 8 ชุดข้อมูล ซึ่งมากกว่าอย่างมีนัยสำคัญ 5 ชุดข้อมูล และความแม่นยำในกลุ่มน้อย วิธี C4.5 – AMOTE มากกว่า วิธี C4.5 – SMOTE จำนวน 8 ชุดข้อมูล โดยมากกว่าอย่างมีนัยสำคัญ 4 ชุดข้อมูล

ตารางที่ 4: สรุปผลค่าความระลึก ค่าความเที่ยง และค่าเอฟ ในชุดข้อมูลกลุ่มน้อย

ชื่อชุดข้อมูล	อัลกอริทึม	ค่า	ค่า	ค่าเอฟ
		ความระลึก	ความเที่ยง	
post-operative ($\alpha = 8$)	C4.5(SMOTE)	0.82	0.69	0.75
	AMOTE	0.83	0.70	0.76
Liver ($\alpha = 2.8$)	C4.5(SMOTE)	0.85	0.71	0.77
	AMOTE	0.85	0.74	0.79
Ionosphere ($\alpha = 5.8$)	C4.5(SMOTE)	0.90	0.88	0.89
	AMOTE	0.92	0.91	0.91
Glasses ($\alpha = 5.7$)	C4.5(SMOTE)	0.93	0.89	0.91
	AMOTE	0.94	0.89	0.92
Ecoli ($\alpha = 6.8$)	C4.5(SMOTE)	0.93	0.95	0.94
	AMOTE	0.93	0.95	0.94
Flags ($\alpha = 3.9$)	C4.5(SMOTE)	0.93	0.93	0.93
	AMOTE	0.94	0.94	0.94
Hepatitis ($\alpha = 2.05$)	C4.5(SMOTE)	0.84	0.87	0.85
	AMOTE	0.85	0.87	0.86
Credit-g ($\alpha = 4$)	C4.5(SMOTE)	0.73	0.73	0.73
	AMOTE	0.75	0.73	0.74

ตารางที่ 5: สรุปผลค่าความระลึก ค่าความเที่ยง และค่าเอฟ ในข้อมูลกลุ่มมาก

ชื่อชุดข้อมูล	อัลกอริทึม	ค่าความระลึก	ค่าความเที่ยง	ค่าเอฟ
post-operative ($\alpha = 8$)	C4.5(SMOTE) AMOTE	0.50 0.50	0.67 0.69	0.57 0.58
Liver ($\alpha = 2.8$)	C4.5(SMOTE) AMOTE	0.50 0.57	0.69 0.72	0.58 0.63
Ionosphere ($\alpha = 5.8$)	C4.5(SMOTE) AMOTE	0.86 0.90	0.89 0.91	0.88 0.90
Glasses ($\alpha = 5.7$)	C4.5(SMOTE) AMOTE	0.32 0.35	0.44 0.52	0.37 0.42
Ecoli ($\alpha = 6.8$)	C4.5(SMOTE) AMOTE	0.79 0.79	0.71 0.73	0.75 0.76
Flags ($\alpha = 3.9$)	C4.5(SMOTE) AMOTE	0.65 0.71	0.63 0.71	0.64 0.71
Hepatitis ($\alpha = 2.05$)	C4.5(SMOTE) AMOTE	0.75 0.75	0.71 0.73	0.73 0.74
Credit-g ($\alpha = 4$)	C4.5(SMOTE) AMOTE	0.69 0.68	0.69 0.70	0.69 0.69

ตารางที่ 6: สรุปผลค่าความระลึก ค่าความเที่ยง และค่าเอฟ ในข้อมูลทั้งหมด

ชื่อชุดข้อมูล	อัลกอริทึม	ค่าความระลึก	ค่าความเที่ยง	ค่าเอฟ
post-operative ($\alpha = 8$)	C4.5(SMOTE) AMOTE	0.68 0.69	0.63 0.63	0.65 0.66
Liver ($\alpha = 2.8$)	C4.5(SMOTE) AMOTE	0.70 0.73	0.70 0.73	0.70 0.73
Ionosphere ($\alpha = 5.8$)	C4.5(SMOTE) AMOTE	0.88 0.91	0.88 0.91	0.88 0.91
Glasses ($\alpha = 5.7$)	C4.5(SMOTE) AMOTE	0.84 0.86	0.84 0.86	0.84 0.86
Ecoli ($\alpha = 6.8$)	C4.5(SMOTE) AMOTE	0.90 0.91	0.90 0.91	0.90 0.91
Flags ($\alpha = 3.9$)	C4.5(SMOTE) AMOTE	0.88 0.91	0.88 0.91	0.88 0.91
Hepatitis ($\alpha = 2.05$)	C4.5(SMOTE) AMOTE	0.81 0.82	0.81 0.82	0.81 0.82
Credit-g ($\alpha = 4$)	C4.5(SMOTE) AMOTE	0.71 0.72	0.71 0.72	0.71 0.72

3.5 สรุปและงานในอนาคต

เมื่อใช้วิธีการ SMOTE เพื่อเพิ่มจำนวนตัวอย่างกลุ่มน้อยจากการสังเคราะห์ขึ้นมาจนมีจำนวนใกล้เคียงกับตัวอย่างกลุ่มมาก เพื่อจำแนกตัวอย่างไม่ให้มีความโน้มเอียงไปในกลุ่มมาก งานวิจัยฉบับนี้จึงนำเสนอการใช้ค่าน้ำหนักต่างกันบนข้อมูลสังเคราะห์หรือข้อมูลกลุ่มน้อย เพื่อจำแนกตัวอย่างกลุ่มน้อยได้ดีขึ้น ซึ่งจากผลการทดลองแสดงให้เห็นว่าการใช้ค่าน้ำหนักต่างกันบนข้อมูลกลุ่มน้อย ที่ระดับต่างๆ (จำนวนเท่า) สำหรับชุดข้อมูลทดสอบแต่ละตัวไม่เท่ากัน เพราะเมื่อทำการเพิ่มจำนวนตัวอย่างกลุ่มน้อยที่ 100% แล้ว ชุดข้อมูลทดสอบบางชุด

จำนวนตัวอย่างกลุ่มน้อยอาจยังมีจำนวนน้อยอยู่เมื่อเทียบกับตัวอย่างกลุ่มมาก ทำให้ผลในการจำแนกตัวอย่างกลุ่มน้อยยังไม่ดี แต่บางชุดข้อมูล เมื่อมีการเพิ่มข้อมูลกลุ่มน้อยจากการสังเคราะห์ขึ้นมาแล้ว ทำให้มีจำนวนกลุ่มน้อยใกล้เคียงกับข้อมูลกลุ่มมาก หรือมากกว่า ทำให้การจำแนกข้อมูลกลุ่มน้อยได้ดีขึ้นเมื่อเทียบกับวิธี C4.5 ที่ใช้เทคนิค SMOTE แสดงให้เห็นว่าเมื่อทำการ SMOTE ที่ 100% จนข้อมูลกลุ่มน้อยมีค่าใกล้เคียงกับกลุ่มมากหรือเท่ากัน จะทำให้การจำแนกข้อมูลกลุ่มน้อยได้ดีขึ้น แต่อาจส่งผลให้การจำแนกตัวอย่างกลุ่มมากอาจลดลงได้ และจากตารางค่าความระลึก ค่าความเที่ยง และค่าเอฟ ในข้อมูลกลุ่มน้อย สรุปได้ว่าวิธีการ AMOTE นั้น มีความแม่นยำเพิ่มขึ้นทั้งหมด 6 ชุดข้อมูล เมื่อเทียบกับวิธี C4.5 ที่ใช้ ชุดข้อมูลที่ทำการ SMOTE เข้าไปทดสอบ แต่ในทางกลับกันเมื่อทำนายในข้อมูลกลุ่มน้อยได้ดี ก็จะไปส่งผลกับข้อมูลกลุ่มมากที่ทำให้มีการจำแนกผิดเพิ่มมากขึ้นด้วย เมื่อใช้ค่าน้ำหนักที่ต่างกันมากขึ้น

ในอนาคตผู้วิจัยวางแผนที่จะพัฒนาส่วนของการจำแนก โดยทำการหาค่าน้ำหนักที่เหมาะสม เมื่อเทียบเป็นสัดส่วนระหว่างข้อมูลกลุ่มมากและข้อมูลกลุ่มน้อย ซึ่งในงานวิจัยนี้ยังเป็นการหาค่าที่ดีที่สุด โดยการทดลอง

เอกสารอ้างอิง

- [1] K. Boonchuay, K. Sinapiromsaran and C. Lursinsap, "Minority Split and Gain Ratio for a Class Imbalance," *Eight International Conference on Fuzzy Systems and Knowledge Discovery (FSKD) 2011*, pp. 2060-3064, 2011.
- [2] อุไรรัตน์ กฤษดาวิชชัย, *ตัวปรับยึดเกาะในดินไม้ตัดสินใจสำหรับเซตข้อมูลไม่สมดุล*, วิทยานิพนธ์, ภาควิชาวิศวกรรมคอมพิวเตอร์, จุฬาลงกรณ์มหาวิทยาลัย, 2553.
- [3] N. V. Chawla, K.W. Bowyer, L. O. Hall, and W. P. Kegelmayer, "SMOTE: Synthetic Minority Over-Sampling Technique," *Journal of Artificial Intelligent Research*, pp. 321-357, 2002.
- [4] ปณต ทรงวัฒนศิริ, *เทคนิคการสุ่มเพิ่มตัวอย่างข้างน้อยสังเคราะห์และเทคนิคการสุ่มลดตัวอย่างข้างมากสำหรับปัญหาความไม่สมดุลระหว่างกลุ่ม*, วิทยานิพนธ์, ภาควิชาคณิตศาสตร์, จุฬาลงกรณ์มหาวิทยาลัย, 2553.
- [5] บุญเสริม กิจศิริกุล, *รายงานการวิจัยฉบับสมบูรณ์ โครงการวิจัยโครงการวิจัยร่วมภาครัฐและเอกชน ปีงบประมาณ 2545 โครงการย่อยที่ 7 อัลกอริทึมการทำเหมืองข้อมูล*, 2546. หน้า: 19-27, 42-44, 82-89.

- [6] บุญเสริม กิจศิริกุล. 2548. ปัญญาประดิษฐ์. กรุงเทพฯ: ภาควิชาวิศวกรรมคอมพิวเตอร์ จุฬาลงกรณ์มหาวิทยาลัย. เวอร์ชัน 1.0.2: 153-162.
- [7] J.R.Quinlan, "Induction of decision trees," *Machine Learning*, Vol.1, pp. 81-106, 1986.
- [8] Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmayer, W. P., "SMOTE: Synthetic Minority Over-Sampling Technique," *Journal of Artificial Intelligent Research* (2002), 321-357.
- [9] C. Bunkhumpornpat, K. Sinapiromsaran and C. Lursinsap, , "Safe-Level-SMOTE: Safe-Level-Synthetic Minority Over-sampling Technique for Handling the Class Imbalanced Problem," *Lecture Notes in Computer Science* (2009), 475-482.
- [10] P. Songwattanasiri, and K. Sinapiromsaran, "Smote: Synthetic Minority Over-sampling and Under-sampling Techniques for class imbalanced problem," *Annual International Conference on computer Science Education Innovation&Technology (CSEIT) 2010*. (2010), 78-83.
- [11] K. Boonchuay, K. Sinapiromsaran and C.Lursinsap, "Minority Split and Gain Ratio for a Class Imbalance," *Eight International Conference on Fuzzy Systems and Knowledge Discovery (FSKD) 2011* (2011), 2060-3064.
- [12] C. L. Blake and C. J. Merz, UCI repository of machine learning databases, Department of Information and Computer Science, [Online], 1998, Available from: <http://www.ics.uci.edu/~mllearn/MLRepository.html>
- [13] Open-Miner Intelligence. (3 พฤศจิกายน 2552). *วิธีการเพิ่มจำนวนข้อมูล (Oversampling) ด้วย SMOTE ใน weka*. [สืบค้นเมื่อ 25 ธันวาคม 2556] จาก http://open-miner.com/2011/06/27/oversampling_smote/
- [14] Weka 3, Data Mining with Open Source Machine Learning Software, [Online], 1999. Available from: <http://www.cs.waikato.ac.nz/~ml/>.

แบบจำลองการพยากรณ์ปริมาณผลผลิตข้าวด้วยเทคนิคโครงข่ายประสาทเทียม Rice Yield Prediction Model based on Artificial Neural Networks

อัทพล คุณเลิศ (Attapol Kunlerd)¹ และบุษยดี ศิริแสงตระกูล (Pusadee Seresangtakul)²

^{1,2}ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น
clashaey@gmail.com¹, pusadee@kku.ac.th²

บทคัดย่อ

บทความนี้มีจุดประสงค์เพื่อศึกษาปัจจัยที่มีผลต่อปริมาณผลผลิตข้าวเพื่อการสร้างตัวแบบการพยากรณ์ปริมาณผลผลิตข้าว ในการศึกษาได้รวบรวมข้อมูลการเพาะปลูกข้าวจากศูนย์วิจัยข้าว 11 ศูนย์ทั่วประเทศผ่านกระบวนการคัดเลือกปัจจัยด้วยวิธีการวิเคราะห์ถดถอยพหุคูณแบบขั้นบันได ผลการศึกษาพบว่ามี 11 ปัจจัยที่มีผลกับปริมาณผลผลิตข้าว ประกอบด้วย พันธุ์ข้าว เดือนที่ปลูก เดือนที่ออกรวง ชนิดดิน อุณหภูมิเฉลี่ยช่วงเดือนที่ 1,3,4,5 ของการปลูกข้าว ปริมาณน้ำฝนเฉลี่ยช่วงเดือนที่ 2,4,5 ของการปลูกข้าว ปัจจัยเหล่านี้ถูกใช้ในการสร้างตัวแบบเพื่อพยากรณ์ผลผลิตข้าวด้วยเทคนิคโครงข่ายประสาทเทียมที่มีกระบวนการเรียนรู้แบบแพร่ย้อนกลับเปรียบเทียบกับวิธีการวิเคราะห์การถดถอยพหุคูณ ผลการทดสอบพบว่า ตัวแบบการพยากรณ์ด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่ย้อนกลับมีประสิทธิภาพในการพยากรณ์สูงกว่าตัวแบบการวิเคราะห์การถดถอยพหุคูณ โดยมีค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ย (RMSE) เท่ากับ 0.0951 และ 0.1273 ตามลำดับ

คำสำคัญ: การพยากรณ์ปริมาณผลผลิตข้าว โครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ การวิเคราะห์การถดถอยพหุคูณ

Abstract

This research aims to study factors affecting rice yield and to presents a model for forecasting rice production. In this study, we collect rice cultivation data from rice research centers (11 centers). In order to look for the appropriate factors, the stepwise regression analysis was

used to analyze the input factors. The result shows that there are 11 factors (rice species, month of rice planting, month of grains, soil type, average temperature during the 1,3,4,5 rice growing month, average rainfall during the 2,4, and 5 rice growing month) have relationship with the rice yield. These factors were used to construct rice yield prediction models by using back propagation artificial neural network (BP-ANN) and multiple regression analysis (MRA) techniques. The prediction results showed that root mean square error (RMSE) of the BP-ANN and MRA models are 0.0951 and 0.1273, respectively. The result showed that the BP-ANN model is effective than the MRA model.

Keywords: Rice Yield Prediction, Back Propagation Neural Network, Multiple Regression Analysis

1. บทนำ

ข้าวเป็นสินค้าเกษตรที่สำคัญของประเทศไทยทั้งในด้านการส่งออกและบริโภค จากรายงานการส่งออกข้าวตั้งแต่เดือนมกราคม - ธันวาคม 2556 ไทยส่งออกข้าวเป็นปริมาณ 6.403 ล้านตัน คิดเป็นมูลค่าการส่งออก 127,008 ล้านบาท [1]

ศูนย์วิจัยข้าว [2] เป็นหน่วยงานส่งเสริมการเพาะปลูกข้าวและทำการทดลองเพาะปลูกข้าว เพื่อเปรียบเทียบปริมาณผลผลิตข้าวและสรุปปัจจัยสภาพแวดล้อมที่เหมาะสม ที่ส่งผลให้ข้าวมีปริมาณผลผลิตที่สูง เพื่อทำการถ่ายทอดเทคโนโลยีและเมล็ดพันธุ์เพื่อการเพาะปลูกแก่เกษตรกร การทดลองปลูกข้าวในแปลงปลูก จำเป็นต้องใช้แรงงานและเวลาในการปลูก ดังนั้นการที่สามารถระบุถึงปัจจัยและปริมาณผลผลิตข้าวที่จะ

ได้รับก่อนการเพาะปลูก จะเป็นแนวทางที่เป็นประโยชน์ ต่อการตัดสินใจของนักวิชาการในด้านการลงทุนเพาะปลูกข้าวในแปลงทดลอง

จากปัญหาที่กล่าวมาในข้างต้น งานวิจัยนี้จึงนำเสนอตัวแบบพยากรณ์เพื่อการคาดการณ์ปริมาณผลผลิตข้าวด้วยเทคนิคโครงข่ายประสาทเทียม [3] ซึ่งมีความสามารถในการเรียนรู้จดจำรูปแบบของชุดข้อมูล สามารถหาความสัมพันธ์ของข้อมูลเพื่อพยากรณ์ข้อมูลที่จะเกิดในอนาคต ในภาคของการเกษตรมีการนำเทคนิคโครงข่ายประสาทเทียมมาใช้ในการพยากรณ์ผลผลิต ตัวอย่างเช่น การพยากรณ์ผลผลิตอ้อย [4], [5] การพยากรณ์ผลผลิตลำไย [6] เป็นต้น

ผลของการวิจัยจะเป็นประโยชน์ต่อการวางแผนและตัดสินใจก่อนการเพาะปลูกข้าว เพื่อลดความเสี่ยงจากการเพาะปลูกข้าวในสภาพแวดล้อมที่ไม่เหมาะสม ต่อการเจริญเติบโตของข้าว

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ปัจจัยที่มีผลต่อผลผลิตข้าว

ข้าว [7] เจริญเติบโตได้ดีตั้งแต่ระดับน้ำทะเลจนถึงความสูง 2500 เมตร เจริญเติบโตได้ดีกับดินทุกชนิดยกเว้นดินทราย มีความต้องการน้ำตั้งแต่ 875 มม. (ข้าวไร่) จนถึง 2000 มม. แสงอาทิตย์มีความจำเป็นในการสังเคราะห์แสง อุณหภูมิมีผลต่อการงอกของเมล็ด การยืดของใบ การแตกกอ และการผสมเกสร เป็นต้น ความชื้นสัมพัทธ์มีผลต่อการพัฒนาของเชื้อโรคของข้าวบางชนิด ลมจะช่วยให้มีการถ่ายเทก๊าซคาร์บอนไดออกไซด์ที่ใช้ในการสังเคราะห์แสง พันธุ์ข้าวแต่ละสายพันธุ์ มีลักษณะและการเจริญเติบโตในสภาพแวดล้อมที่แตกต่างกัน [8] ระยะเวลาที่ใช้ในการเพาะปลูกข้าวจะส่งผลกระทบต่อปริมาณน้ำของต้นข้าว [9] ปริมาณน้ำมีอิทธิพลต่อผลผลิตข้าว [10]

2.2 โครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ (Back Propagation Neural Network)

โครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ [3] ประกอบด้วย 3 ชั้น ได้แก่ ชั้นอินพุต ชั้นซ่อน และชั้นเอาต์พุต โดยขั้นตอนวิธีการเรียนรู้จะแบ่งออกเป็นสองขั้นตอน คือ ขั้นตอนการป้อนไปข้างหน้าและขั้นตอนการแพร่ย้อนกลับ กระบวนการทำงานจะเริ่มจากการป้อนข้อมูลในชั้นอินพุต

คำนวณค่าเอาต์พุตในแต่ละชั้นไปจนถึงชั้นเอาต์พุต นำผลลัพธ์ของข้างงานหาค่าความผิดพลาด ด้วยการคำนวณผลต่างของผลลัพธ์กับค่าเป้าหมายโดยใช้สมการที่ (1) ในการคำนวณ จากนั้นจะเป็นกระบวนการทำงานแบบแพร่ย้อนกลับ โดยนำค่าผิดพลาดที่คำนวณได้จากสมการที่ (1) ไปปรับค่าน้ำหนักในแต่ละชั้น วนซ้ำการทำงานจนกว่าค่าความผิดพลาดที่ได้จะอยู่ระดับที่กำหนดไว้

$$E(t) = \sum_{k=1}^K \sum_{i=1}^{N(L)} (d_i(k) - z_i(k))^2 \quad (1)$$

โดย K แทน จำนวนข้อมูลฝึกสอนต่อรอบการทำงาน
 $N(L)$ แทน จำนวนโหนดในชั้นที่ L
 $d_i(k)$ แทน ค่าข้อมูลจริงของโหนดที่ i
 $z_i(k)$ แทน ค่าเอาต์พุตของโหนดที่ i
 $E(t)$ แทน ค่าผลรวมค่าความแตกต่างระหว่างค่าเป้าหมายและค่าข้อมูลจริง โดยเป็นผลรวมของทุกโหนดในชั้นเอาต์พุตและทุกข้อมูลฝึกสอน ณ การทำงานรอบที่ t

2.3 การวิเคราะห์การถดถอยพหุคูณ

การวิเคราะห์การถดถอยพหุคูณ [11] เป็นการวิเคราะห์ข้อมูลเพื่อหาความสัมพันธ์ระหว่างตัวแปรตาม (Y) จำนวน 1 ตัว กับตัวแปรอิสระ (X) ตั้งแต่ 2 ตัวขึ้นไป อาศัยความสัมพันธ์เชิงเส้นตรงระหว่างตัวแปรเพื่อใช้ในการทำนาย ผลของการวิเคราะห์ข้อมูลจะอยู่ในรูปของสมการเชิงเส้นตรงในรูปคะแนนดิบ ดังสมการที่ (2)

$$Y' = a + b_1x_1 + b_2x_2 + \dots + b_kx_k \quad (2)$$

โดยที่ Y' แทน คะแนนพยากรณ์ตัวแปรตาม (Y)
 a แทน ค่าคงที่ของสมการ
 b แทน สัมประสิทธิ์การถดถอยของตัวแปรอิสระ
 x แทน คะแนนตัวแปรอิสระ
 k แทน จำนวนตัวแปรอิสระ (X)

2.4 การตรวจสอบแบบไขว้ (k-fold Cross-validation)

การตรวจสอบแบบไขว้ [12] เป็นวิธีการทดสอบโดยแบ่งกลุ่มข้อมูลออกเป็น k กลุ่ม โดยข้อมูลแต่ละกลุ่มจะมีจำนวนเท่ากัน ข้อมูลแต่ละกลุ่มจะถูกวนให้เป็นข้อมูลการสอนและข้อมูลทดสอบ วิธีนี้จะแก้ปัญหาการเลือกข้อมูลที่ดีและง่ายมาเป็นข้อมูลทดสอบ ในแต่ละรอบข้อมูลจำนวน $k-1$ จะถูกใช้

เป็นข้อมูลในการสอนและข้อมูล 1 กลุ่มจะถูกใช้เป็นข้อมูลในการทดสอบ จะทำการหมุนเวียนไปเรื่อยๆจนครบ k กลุ่ม จากนั้นหาค่าเฉลี่ยความถูกต้องจากทั้ง k รอบ ตัวอย่างเช่น การแบ่งชุดข้อมูลออกเป็น 5 กลุ่ม หรือ k = 5 โดยในรอบที่ 1 จะใช้ชุดข้อมูล 2-5 เป็นชุดข้อมูลการสอนและใช้ข้อมูลชุดที่ 1 เป็นชุดข้อมูลทดสอบ และทำการสับเปลี่ยนข้อมูลจนครบ 5 รอบ อธิบายการทำงานดังแสดงในภาพที่ 1

Iteration 1: train on	2 3 4 5	test on	1
Iteration 2: train on	1 3 4 5	test on	2
Iteration 3: train on	1 2 4 5	test on	3
Iteration 4: train on	1 2 3 5	test on	4
Iteration 5: train on	1 2 3 4	test on	5

ภาพที่ 1: แสดงการแบ่งชุดข้อมูลแบบ k-fold โดยที่ k=5 [12]

2.4 การวัดประสิทธิภาพการพยากรณ์

งานวิจัยนี้ทำการเปรียบเทียบประสิทธิภาพการพยากรณ์ของ 2 ตัวแบบ โดยใช้ตัวชี้วัดคือ รากที่สองของความคลาดเคลื่อนเฉลี่ย (root mean square error: RMSE) [13] คำนวณได้จากสมการที่ (3) ซึ่งเป็นการวัดความแตกต่างระหว่างค่าจริงและค่าที่ได้จากการพยากรณ์ โดยค่าที่เข้าใกล้ 0 แสดงว่าตัวแบบมีความแม่นยำในการพยากรณ์สูง

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y'_i - Y_i)^2} \tag{3}$$

โดย Y_i แทน ค่าของข้อมูลที่เกิดขึ้นจริง
 Y'_i แทน ค่าที่ได้จากการพยากรณ์
 n แทน จำนวนนำเข้าทั้งหมด

2.5 งานวิจัยที่เกี่ยวข้อง

รัชภูมิ ใจกล้า [14] นำเสนอการทำนายผลผลิตข้าวด้วยซอฟต์แวร์เวกเตอร์แมชชีน แบ่งการทำนายออกเป็น 3 ขั้นตอน ได้แก่ ทำนายน้ำหนักรวมในไตรเจน น้ำหนักต้นข้าว และน้ำหนักเมล็ดข้าว โดยเปรียบเทียบการทดลองกับการวิเคราะห์การถดถอยเชิงเส้นและไม่เชิงเส้น ใช้ข้าวพันธุ์ กข.23 ปัจจัยนำเข้าประกอบด้วย ข้อมูลรังสีดวงอาทิตย์ อุณหภูมิสูงสุด-ต่ำสุด ปริมาณน้ำฝน ปริมาณน้ำในนาข้าว ผลการทดลองแสดงให้เห็นว่าวิธีการซอฟต์แวร์เวกเตอร์แมชชีนเป็นตัวแบบที่มีประสิทธิภาพสูงสุด สามารถทำนายน้ำหนักเมล็ดข้าว ให้ค่าเปอร์เซ็นต์ผิดพลาดสัมบูรณ์เฉลี่ยเท่ากับ 3.74 และการวิเคราะห์

ถดถอยแบบเชิงเส้นและไม่เชิงเส้นมีผลการทดลองเท่ากับ 24.30 และ 11.87 ตามลำดับ

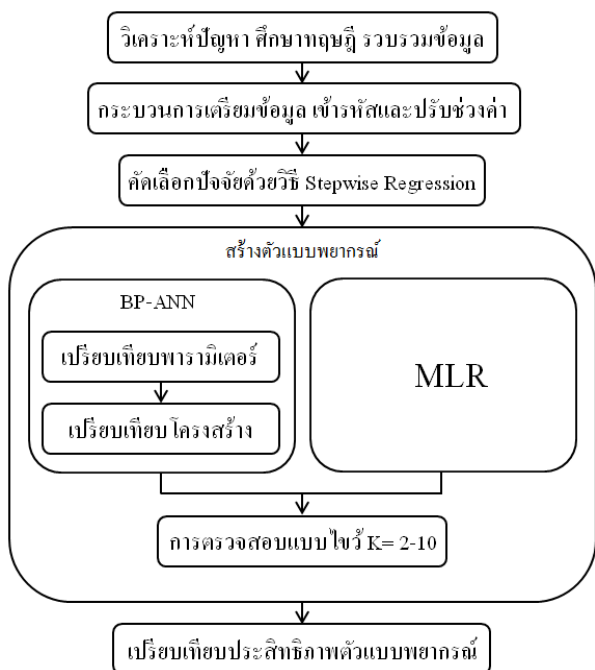
Ibrahim และ O.M. [15] สร้างตัวแบบการพยากรณ์ปริมาณผลผลิตข้าวบาร์เลย์ โดยเปรียบเทียบโครงข่ายประสาทเทียมและการวิเคราะห์การถดถอยเชิงเส้นแบบพหุ ในส่วนของตัวแบบโครงข่ายประสาทเทียมสร้าง 2 ตัวแบบ คือ 1.General regression neural network 2.Multilayer perceptron ปัจจัยนำเข้าตัวแบบประกอบด้วย ความสูงของข้าว ความยาวของรวงข้าว น้ำหนักเมล็ดข้าว ฟางข้าวและลักษณะทางชีวภาพ ผลการศึกษาพบว่า General regression neural network มีความแม่นยำสูงกว่าการวิเคราะห์การถดถอยเชิงเส้นแบบพหุ โดยมีค่าสัมประสิทธิ์สหสัมพันธ์ 93.4 % และ 86.6 % ตามลำดับ

Zhiliang Wang, Binbin Li และ Lei Cao [16] นำเสนอการพยากรณ์ผลผลิตข้าวด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ โดยเปรียบเทียบประสิทธิภาพของตัวแบบที่มีการวิเคราะห์ปัจจัย ด้วยกระบวนการวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis: PCA) และการนำปัจจัยทั้งหมดเข้าสู่ตัวแบบ ประกอบด้วย จำนวนแรงงาน พื้นที่ทำการเกษตร จำนวนเครื่องจักร ปริมาณน้ำ ปริมาณของปุ๋ยที่ใช้ การใช้พลังงานในพื้นที่ สารกำจัดศัตรูพืชทั้งหมด และปริมาณสารที่ใช้แต่ละครั้ง ผลการทดลองพบว่า วิธีการวิเคราะห์ปัจจัยด้วยการวิเคราะห์องค์ประกอบหลัก ให้ประสิทธิภาพสูงกว่าในแบบที่ไม่มีการวิเคราะห์ปัจจัย โดยให้ค่าความผิดพลาดต่ำสุดที่ 0.57 % และ 1.24% ตามลำดับ

Mohammad Zaefizadeh, Majid Khayatnezhad และ Roza gholamin [17] สร้างตัวแบบเพื่อทำการพยากรณ์ผลผลิตข้าวบาร์เลย์ 40 สายพันธุ์ โดยทำการเปรียบเทียบตัวแบบระหว่าง โครงข่ายประสาทเทียมและการวิเคราะห์การถดถอยเชิงเส้นแบบพหุ โดยใช้ข้อมูลการปลูกข้าว จากแปลงทดลองของมหาวิทยาลัย Islamic Azad ในเมือง Ardabil ประเทศอิหร่าน ซึ่งใช้ 3 ปัจจัยในการวิเคราะห์ประกอบด้วย จำนวนของต้นกล้า จำนวนเมล็ดข้าวต่อรวงและน้ำหนักของเมล็ดข้าว ผลการวิจัยแสดงให้เห็นว่าตัวแบบโครงข่ายประสาทเทียมมีประสิทธิภาพสูงกว่าตัวแบบการวิเคราะห์การถดถอยเชิงเส้นแบบพหุ โดยให้ค่าเบี่ยงเบนมาตรฐานที่ 0.597 และ 1.074 ตามลำดับ

3. วิธีการดำเนินงาน

งานวิจัยนี้มีขั้นตอนในการดำเนินการวิจัย ดังภาพที่ 2



ภาพที่ 2: ขั้นตอนการดำเนินงานวิจัย

3.1 การรวบรวมและจัดเตรียมข้อมูล

ในการศึกษานี้ รวบรวมข้อมูลการเพาะปลูกข้าวระหว่าง พ.ศ. 2551-2555 จากจากศูนย์วิจัยข้าวทั่วประเทศจำนวน 11 ศูนย์ ได้แก่ ศูนย์วิจัยข้าวจังหวัดขอนแก่น แม่ฮ่องสอน เชียงใหม่ เชียงราย แพร่ นongคาย สกลนคร อุดรธานี นครราชสีมา อุบลราชธานีและจังหวัดสุรินทร์ จำนวน 2404 ไร่เบียน รวมถึงข้อมูลสภาพอากาศจากกรมอุตุนิยมวิทยา และ ข้อมูลชนิดดินของพื้นที่ทำการเพาะปลูกจากเว็บไซต์กรมพัฒนาที่ดิน จำนวน 17 ปัจจัย ดังแสดงรายละเอียดในตารางที่ 1

เมื่อทำการรวบรวมปัจจัยทั้งหมดแล้ว จากนั้นทำการตรวจสอบความถูกต้องของข้อมูล เนื่องจากข้อมูลบางส่วนไม่สมบูรณ์ ดังนั้นในขั้นตอนของการทำความสะอาดข้อมูล (Data cleaning) งานวิจัยนี้เลือกจัดการข้อมูลสูญหายด้วย วิธีการตัดแถวข้อมูลที่ไม่มีสมบูรณ์ทั้ง (Listwise deletion) [18] จากนั้นทำการเข้ารหัสข้อมูลและทำการแปลงข้อมูลให้อยู่ในช่วง 0-1 ด้วยวิธีการ Min-max normalization [19] คำนวณจากสมการที่ (4)

$$Out = \frac{Input - Min}{Max - Min} \tag{4}$$

โดยที่ *Input* แทน ข้อมูลที่นำมาคำนวณในแต่ละตำแหน่ง
Min แทน ค่าที่น้อยที่สุดของชุดข้อมูลนั้น
Max แทน ค่าที่มากที่สุดของชุดข้อมูลนั้น

ตารางที่ 1: แสดงปัจจัยและการจัดเก็บข้อมูล

ปัจจัย	ข้อมูล
ปีที่ปลูก (Year)	2551
พันธุ์ข้าว (RiceType)	9 = พันธุ์ 603 54R
เดือนที่ปลูก (GrowingMonth)	7 = กรกฎาคม
เดือนออกรวง (FlowerDate)	10 = ตุลาคม
จำนวนวันที่ออกรวง (DurFlower)	130
ชนิดดิน (SoilType)	1 = 22/24
อุณหภูมิเฉลี่ยช่วงเดือนที่ 1 (M1AT)	ค่าตัวเลขอุณหภูมิ เช่น 27.6
อุณหภูมิเฉลี่ยช่วงเดือนที่ 2 (M2AT)	ค่าตัวเลขอุณหภูมิ เช่น 27.6
อุณหภูมิเฉลี่ยช่วงเดือนที่ 3 (M3AT)	ค่าตัวเลขอุณหภูมิ เช่น 27.3
อุณหภูมิเฉลี่ยช่วงเดือนที่ 4 (M4AT)	ค่าตัวเลขอุณหภูมิ เช่น 26.9
อุณหภูมิเฉลี่ยช่วงเดือนที่ 5 (M5AT)	ค่าตัวเลขอุณหภูมิ เช่น 24.2
ฝนเฉลี่ยช่วงเดือน 1 (M1AR)	ค่าปริมาณน้ำฝนเช่น 185.4
ฝนเฉลี่ยช่วงเดือน 2 (M2AR)	ค่าปริมาณน้ำฝน เช่น 179.6
ฝนเฉลี่ยช่วงเดือน 3 (M3AR)	ค่าปริมาณน้ำฝน เช่น 80.1
ฝนเฉลี่ยช่วงเดือน 4 (M4AR)	ค่าปริมาณน้ำฝน เช่น 38.8
ฝนเฉลี่ยช่วงเดือน 5 (M5AR)	ค่าตัวเลขปริมาณน้ำฝน เช่น 1
ผลผลิต	ค่าตัวเลขผลผลิต

3.2 วิเคราะห์ความสัมพันธ์ของปัจจัย

จากปัจจัยที่รวบรวมไว้ในตารางที่ 1 ข้างต้น ทำการวิเคราะห์ปัจจัย ด้วยวิธีการวิเคราะห์ถดถอยพหุคูณแบบขั้นบันได (Stepwise regression) [20] ซึ่งเป็นเทคนิคของการลดตัวแปร [21] ผลการวิเคราะห์ พบว่าปัจจัยที่ส่งผลกับปริมาณผลผลิตข้าวมีจำนวนทั้งสิ้น 11 ปัจจัย ประกอบด้วย พันธุ์ข้าว เดือนที่ปลูก เดือนที่ออกรวง ชนิดดิน อุณหภูมิเฉลี่ยช่วงเดือนที่ 1,3,4,5 และปริมาณฝนเฉลี่ยช่วงเดือน 2, 4 และ 5 ของการเพาะปลูก

3.3 ตัวแบบเพื่อพยากรณ์ผลผลิตข้าว

งานวิจัยนี้ทำการสร้างตัวแบบโดยใช้เทคนิคโครงข่ายประสาทเทียมแบบแพร่ย้อนกลับและการวิเคราะห์การถดถอยพหุคูณ โดยใช้ปัจจัยนำเข้า 11 ปัจจัยที่ได้จากการคัดเลือกปัจจัย

3.3.1 ตัวแบบโครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ

การสร้างตัวแบบโครงข่ายประสาทเทียมแบบแพร่ย้อนกลับเพื่อให้ได้ตัวแบบที่มีความเหมาะสม จะทำการปรับค่าพารามิเตอร์ ซึ่งประกอบด้วย อัตราการเรียนรู้ 0.01- 0.5 โมเมนตัม 0.01- 0.5 และจำนวนรอบการเรียนรู้ 1000-5000 ผลการศึกษาสามารถสรุปค่าพารามิเตอร์ที่ให้ผลของความผิดพลาดในการพยากรณ์น้อยที่สุดแสดงในตารางที่ 2

ตารางที่ 2: ค่าพารามิเตอร์ที่ดีที่สุดในการพยากรณ์

พารามิเตอร์	ค่า
จำนวน Input	11
จำนวน Output	1
จำนวนโหนดชั้นซ่อน	10 , 2
อัตราการเรียนรู้	0.02
โมเมนตัม	0.02
จำนวนรอบการเรียนรู้	5000

เมื่อได้ค่าพารามิเตอร์ที่ดีที่สุดแล้ว จะทำการเปรียบเทียบโครงสร้างของโครงข่ายประสาทเทียม เพื่อหาโครงสร้างที่เหมาะสม โดยการแบ่งชุดข้อมูลการสอนและการทดสอบด้วยการตรวจสอบแบบไขว้ ซึ่งเปรียบเทียบค่า k ตั้งแต่ 2 k - 10 k และวัดประสิทธิภาพการพยากรณ์ด้วยค่ารากที่สองของความคลาดเคลื่อนเฉลี่ย

3.3.2 ตัวแบบการวิเคราะห์การถดถอย

ในการพัฒนาตัวแบบการวิเคราะห์การถดถอยพหุคูณ ใช้ปัจจัยนำเข้าที่ได้จากขั้นการวิเคราะห์ปัจจัย ตัวแปรที่ต้องการพยากรณ์คือปริมาณผลผลิตข้าว โดยแบ่งชุดข้อมูลด้วยการทดสอบแบบไขว้กำหนดค่า k=2-10 ซึ่งสมการพยากรณ์ที่ได้แสดงในสมการที่ (5)

$$\begin{aligned}
 \text{ผลผลิตต่อไร่} = & 0.578 + 0.0117(\text{RiceType}) \\
 & + 0.1132(\text{GrowingMonth}) + 0.1522(\text{FlowerDate}) \\
 & + 0.0855(\text{SoilType}) + 0.1508(\text{MIAT}) + 0.2119(\text{M3AT}) \\
 & + 0.0858(\text{M4AT}) + 0.31(\text{M5AT}) + 0.0426(\text{M2AR}) \\
 & + 0.0496(\text{M4AR}) + 0.0704(\text{M5AR}) \quad (5)
 \end{aligned}$$

4. ผลการดำเนินการวิจัย

จากการทดลองเปรียบเทียบเพื่อหาความแม่นยำในการพยากรณ์ โดยใช้เทคนิคโครงข่ายประสาทเทียมแบบแพร่ย้อนกลับและการวิเคราะห์การถดถอยพหุคูณ โดยการแบ่งชุดข้อมูลการตรวจสอบแบบไขว้ ค่า k ตั้งแต่ 2 k - 10 k และวัดประสิทธิภาพการพยากรณ์ด้วยค่ารากที่สองของความคลาดเคลื่อนเฉลี่ย ผลการทดลองแสดงในตารางที่ 3

ตารางที่ 3: ผลการเปรียบเทียบประสิทธิภาพการพยากรณ์ด้วยเทคนิคโครงข่ายประสาทเทียมและการวิเคราะห์การถดถอย

k-fold	การวิเคราะห์การถดถอย (MLR)	โครงข่ายประสาทเทียม (BP-ANN)
	RMSE	RMSE
2 fold	0.1279	0.0979
3 fold	0.1279	0.0980
4 fold	0.1275	0.0975
5 fold	0.1276	0.0970
6 fold	0.1275	0.0961
7 fold	0.1274	0.0954
8 fold	0.1274	0.0963
9 fold	0.1273	0.0951
10 fold	0.1274	0.0966

จากข้อมูลในตารางที่ 3 เมื่อทำการเปรียบเทียบประสิทธิภาพของการพยากรณ์ของทั้ง 2 ตัวแบบพบว่า เทคนิคโครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ มีค่ารากที่สองของความคลาดเคลื่อนเฉลี่ยต่ำสุดที่ 0.951 (ที่ 9 fold) ซึ่งมีประสิทธิภาพดีกว่าวิธีการวิเคราะห์การถดถอยพหุคูณที่มีค่ารากที่สองของความคลาดเคลื่อนเฉลี่ยต่ำสุดที่ 0.1273 (ที่ 9 fold)

5. สรุปผลการวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อสร้างตัวแบบเพื่อการพยากรณ์ปริมาณผลผลิตข้าว โดยใช้เทคนิคโครงข่ายประสาทเทียมแบบแพร่ย้อนกลับและการวิเคราะห์การถดถอยพหุคูณ ผลของการทดลองพบว่า ตัวแบบเทคนิคโครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ มีประสิทธิภาพในการพยากรณ์สูงกว่าตัวแบบการ

วิเคราะห์การถดถอย โดยมีค่ารากที่สองของความคลาดเคลื่อนเฉลี่ยเท่ากับ 0.0951 และ 0.1273 ตามลำดับ

แนวทางในการศึกษาต่อไปควรมุ่งศึกษาปัจจัยด้านการเกษตรที่มีผลต่อผลผลิตข้าวเพิ่มเติม เช่น ปริมาณแสงแดด การระบาดของโรค เป็นต้น

เอกสารอ้างอิง

- [1] สมาคมผู้ส่งออกข้าว, “สถิติส่งออกข้าว,” [ออนไลน์] 2556. สืบค้นจาก: http://www.thairiceexporters.or.th/default_th.htm [5 พฤศจิกายน 2556].
- [2] สำนักวิจัยและพัฒนาข้าว, “การกิจและหน้าที่,” [ออนไลน์] 2556. สืบค้นจาก: <http://www.brrd.in.th/main/> [5 พฤศจิกายน 2556].
- [3] Simon Haykin, *Neural Networks - A Comprehensive Foundation*, Pearson education: India, 1998.
- [4] อรณรงค์ บุญเกิด และ พุทธิศิรี สิริแสงตระกูล, “แบบจำลองการพยากรณ์ปริมาณอ้อยของภาคตะวันออกเฉียงเหนือโดยการใช้โครงข่ายประสาทเทียม,” *วารสารวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น*, ปีที่ 41 ฉบับที่ 1 มกราคม - มีนาคม พ.ศ. 2556 หน้า 213-255.
- [5] สุรเชษฐ์ เกื้อนแก้วสิงห์, “การใช้โครงข่ายประสาทเทียมและระบบสารสนเทศภูมิศาสตร์ในการพยากรณ์ผลผลิตอ้อย,” *วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมอุตสาหการ บัณฑิตวิทยาลัย มหาวิทยาลัยขอนแก่น*, 2552.
- [6] กาญจนา ทองบุญนาค, “การบูรณาการโครงข่ายประสาทเทียมและระบบสารสนเทศภูมิศาสตร์เพื่อการพยากรณ์ผลผลิตทางการเกษตร: กรณีศึกษาผลผลิตลำไยในพื้นที่จังหวัดเชียงใหม่และลำพูนประเทศไทย,” *วิทยานิพนธ์ปริญญาวิทยาศาสตรดุษฎีบัณฑิต สาขาวิชาภูมิสารสนเทศ บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีสุรนารี*, 2551.
- [7] คณะทรัพยากรธรรมชาติมหาวิทยาลัยสงขลานครินทร์, “ข้าว,” [ออนไลน์] 2556 สืบค้นจาก: <http://www.natres.psu.ac.th/Department/PlantScience/510-211/lecturenote/document/rice.doc> [5 พฤศจิกายน 2556].
- [8] นิตยา รื่นสุข, ประนอม มงคลบรรจบ, สุภาพร จันทร์บัวทอง และ วาสนา อินแถลง, “การเปรียบเทียบผลผลิตข้าวลูกผสม ข้าวปทุมธานี 1 และข้าวพิษณุโลก 2 ภายใต้การปลูกโดยวิธีหว่านน้ำตมโดยใช้อัตราเมล็ดพันธุ์ 4 กก./ไร่,” *ผลงานวิจัยด้านการเกษตรกรรม ศูนย์วิจัยข้าวปทุมธานี กรมการข้าว สำนักวิจัยและพัฒนาข้าว*, 2550, หน้า 18-20.
- [9] สุรจิต ภูภักดี และบาส เบาว์แมนน์, “การวิเคราะห์ความเสี่ยงในการปลูกข้าวน้ำฝนโดยใช้ระบบสารสนเทศภูมิศาสตร์และแบบจำลองการเจริญเติบโตของพืช,” *วารสารวิชาการข้าว* ปีที่ 2 ฉบับที่ 1 มกราคม-เมษายน พ.ศ. 2551 หน้า 13-25.
- [10] สำนักวิจัยและพัฒนาข้าว, “การใช้ปุ๋ยน้ำอย่างประหยัด,” [ออนไลน์] 2556. จาก: http://www.brrd.in.th/rkb/Fact%20Sheet/Fertilizer/Fertilizer_013.pdf [5 พฤศจิกายน 2556].
- [11] Lenona S. Aiken and Stephen G. West, “*Multiple regression: Testing and interpreting interactions*,” Sage publications: California, 1991.
- [12] Ron Kohavi, “A study of cross-validation and bootstrap for accuracy estimation and model selection,” *Proceeding IJCAI/95 Proceedings of the 14th International Joint conference on Artificial intelligence*, pp. 1137-1143, 20-25 August 1995.
- [13] ทรงศิริ แด่สมบัติ, “เทคนิคการพยากรณ์เชิงปริมาณ,” *ฟิสิกส์เชิงเตอร์: กรุงเทพฯ*, 2539.
- [14] รัชภูมิ ใจกล้า, “การทำนายผลผลิตข้าวโดยใช้ระบบการอนุมานฟuzzy และซัพพอร์ตเวกเตอร์แมชชีน,” *วิทยานิพนธ์วิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมคอมพิวเตอร์ บัณฑิตวิทยาลัย มหาวิทยาลัยเชียงใหม่*, 2552.
- [15] Hassan O.M.I., “Simulation of Barley grain yield using artificial neural networks and multiple linear regression models,” *Egyptian Journal of Applied Sciences*, vol.27, no.1, pp.1-11, 2012.
- [16] Wang Z., Li B. and Cao L., “Application of BP Neural Network Based on Principal Component Analysis in Grain Yield Prediction,” *Information Science and Engineering (ICISE)*, Dec 4-6, 2012, pp.1011 – 1013.
- [17] Zaefizadeh M., Khayatnezhad M. and Gholamin R., “Comparison of Multiple Linear Regressions (MLR) and Artificial Neural Network (ANN) in Predicting the Yield Using its Components in the Hulled Barley,” *American-Eurasian journal of agricultural and environmental sciences*, vol.10, no.1, pp.60-64, 2011.
- [18] Little R.J. and Rubin D.B., “*Statistical Analysis with Missing Data*,” John Wiley and Sons: New York, 1987.
- [19] R. Begg and M. Palaniswami, “Computational Intelligence for Movement Sciences,” Idea group: London, 2006.
- [20] Harrell, Frank E., *Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis*, 2nd Edition, Springer-Verlag: New York, 2001.
- [21] Pádraig Cunningham, “Dimension Reduction,” *Machine Learning Techniques for Multimedia, Cognitive Technologies*, 2008, pp.91-112.

การหาค่าถ่วงน้ำหนักสำหรับเคมีนโดยใช้ขั้นตอนวิธีเชิงพันธุกรรม

Genetic Algorithm for Weighted K-Means

ณัฐวุฒิ ปั้นรูป (Natthawoot Punroob)¹ อัฐพร กิ่งบุญ (Athaporn Kingboo)² และ Guangqian Chen³

^{1,3} คณะบริหารธุรกิจและศิลปศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา ตาก

² คณะศิลปศาสตร์และวิทยาศาสตร์ มหาวิทยาลัยราชภัฏศรีสะเกษ

natthakub2inter@outlook.com¹, atha@sskru.ac.th², yingyichen@outlook.com³

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์นำเสนอวิธีการจำแนกข้อมูลด้วยวิธีการจัดกลุ่มข้อมูล โดยการแบ่งข้อมูลออกตามคุณลักษณะข้อมูล ในการจัดกลุ่มข้อมูลได้นำเทคนิคจัดกลุ่มเคมีน ซึ่งมีกำหนดค่าน้ำหนักคุณลักษณะที่เหมาะสมและจำนวนกลุ่มที่เหมาะสมด้วยขั้นตอนวิธีเชิงพันธุกรรม (Genetic Algorithm :GA) เป็นการเพิ่มประสิทธิภาพด้านความถูกต้องให้สูงขึ้น ผู้วิจัยได้ออกแบบโมเดลการหาค่าถ่วงน้ำหนักสำหรับเคมีนโดยใช้ขั้นตอนวิธีเชิงพันธุกรรม (Genetic Algorithm for Weighted K-means: GAWK) จากการวิจัยการจัดกลุ่มข้อมูลภาวะการมีงานทำของบัณฑิตมหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา ตาก โดยโมเดล GAWK พบว่าการแบ่งกลุ่มตามระดับเงินเดือนด้วยเทคนิคจัดกลุ่มเคมีนมีค่าความถูกต้อง 74.99 % และหลังจากใช้โมเดล GAWK มีค่าความถูกต้อง 85.24 % แสดงให้เห็นว่าถ้ามีการกำหนดค่าน้ำหนักที่เหมาะสมให้กับการจัดกลุ่มเคมีน ทำให้การจัดกลุ่มข้อมูลมีความถูกต้องมากยิ่งขึ้น

คำสำคัญ: เหมือนข้อมูล การจัดกลุ่ม เคมีน น้ำหนัก ขั้นตอนวิธีเชิงพันธุกรรม

Abstract

This research aims to present how to make classification from various data. A model which is combined K-Means, Weighted and Genetic Algorithms is created for increasing the accuracy of classification, named GAWK (Genetic Algorithm for Weighted K-means). The data is job status of graduated students of Rajamangala University of Technology Lanna, Tak. The findings show that comparing the classification result from K-Means, which is 74.99%, the classification result

is 85.24% using GAWK method. That means GAWK method can increase the accuracy of data classification.

Keyword: Data Mining, Clustering, K-Means, Weighted, Genetic Algorithm

1. บทนำ

ในปัจจุบันเป็นยุคข้อมูลสารสนเทศในลักษณะการให้บริการคำปรึกษาเพื่อช่วยในการตัดสินใจและเทคโนโลยีอินเทอร์เน็ตถูกนิยมมาประยุกต์เพื่อให้บริการคำปรึกษาการเลือกสมัครงานในตำแหน่งที่เหมาะสมแก่บัณฑิตจบใหม่มีจำนวนมาก การให้บริการคำปรึกษาที่แม่นยำนั้นสำคัญมากแต่การนำข้อมูลที่เกี่ยวข้องบนฐานข้อมูลทั้งหมดมาใช้สืบค้นเพื่อให้บริการคำปรึกษาแต่ละครั้งจะใช้เวลาในการประมวลผลค่อนข้างนาน งานวิจัยการทำเหมืองข้อมูล (Data Mining) เพื่อสกัดข้อมูลสารสนเทศบนเว็บไซต์ด้วยวิธีการจัดกลุ่มข้อมูล (Clustering) ซึ่งเทคนิคที่นิยมใช้กับการบริการคำปรึกษาบนเว็บไซต์คือเทคนิคจัดกลุ่มข้อมูลเคมีน (K-means Clustering) ก็จะนำข้อมูลที่มีลักษณะคล้ายกันจัดกลุ่มเดียวกันและตัวแทนข้อมูลของทุกกลุ่ม (Centroid) คือตัวแทนข้อมูลสมาชิกทั้งหมดทำให้ได้ผลลัพธ์การค้นหาข้อมูลบนเว็บไซต์มีความรวดเร็ว [1] และการจัดกลุ่มข้อมูลไม่ต้องการมีผู้เชี่ยวชาญช่วยสอน (Unsupervised Learning) [2] แต่ข้อจำกัดการใช้เทคนิคจัดกลุ่มเคมีนคือผู้ใช้งานจำเป็นต้องรู้จำนวนกลุ่มที่จะเหมาะสม ก่อนนำข้อมูลไปวิเคราะห์ ในปัจจุบันเทคนิคการจัดกลุ่มเคมีนได้รับความนิยมและมีการใช้ในซอฟต์แวร์ประยุกต์ทางด้านการทำเหมืองข้อมูล เช่น MATLAB, SPSS และ WAKA เพื่อใช้ในการวิเคราะห์ข้อมูล [3] การนำเทคนิค GA เป็นเทคนิคเพื่อค้นหาประสิทธิภาพที่ดีที่สุดได้นำมาทำงานร่วมกับเทคนิคจัดกลุ่มเคมีน โดยเทคนิค GA จะค้นหาเฉพาะจำนวนกลุ่มที่เหมาะสมให้กับ

เทคนิคจัดกลุ่มเคมีนซึ่งได้ผลลัพธ์ประสิทธิภาพด้านความถูกต้องดีขึ้นเมื่อเปรียบเทียบกับการใช้เทคนิคจัดกลุ่มเคมีนเพียงเทคนิคเดียว [4][5] และได้ผลลัพธ์ประสิทธิภาพในการใช้เวลาคำนวณได้เร็วขึ้นเมื่อเปรียบเทียบกับการใช้เทคนิค GA เพียงเทคนิคเดียว[6] แต่ข้อจำกัดการทำงานร่วมกันทั้งสองเทคนิคคือ ลักษณะข้อมูลทั้งหมดจะถูกกำหนดความสำคัญเท่ากัน [4][5] [6] หากมีลักษณะข้อมูลด้านใดเป็นปัจจัยที่มีความสำคัญต่อผลลัพธ์ไม่เท่ากันประสิทธิภาพการค้นหาคำตอบที่แม่นยำมีค่าน้อยลง โดยผู้วิจัยเห็นควรมีการเพิ่มน้ำหนักความสำคัญให้กับคุณลักษณะข้อมูลซึ่งเป็นการเพิ่มประสิทธิภาพเมื่อนำสองเทคนิคทำงานร่วมกัน

การศึกษาถึงการกำหนดน้ำหนัก (Weighted) ให้กับคุณลักษณะข้อมูลที่มีความสำคัญไม่เท่ากัน ด้วยเทคนิคการจัดกลุ่มมินโคสกี (Minkowski Weighted K-means) [3] ซึ่งทำให้ประสิทธิภาพการค้นหาคำตอบที่แม่นยำมีค่ามากขึ้นมีค่ามากขึ้น แต่ก็ยังคงมีข้อจำกัดที่ต้องกำหนดจำนวนกลุ่มก่อนการวิเคราะห์การจัดกลุ่มข้อมูล จากการศึกษาที่กล่าวมาข้างต้นผู้วิจัยจึงประยุกต์นำเทคนิค GA และเทคนิคจัดกลุ่มเคมีนทำงานร่วมกัน มีชื่อเรียกโมเดลนี้ว่าการหาค่าถ่วงน้ำหนักสำหรับเคมีน โดยใช้ขั้นตอนวิธีเชิงพันธุกรรม (Genetic Algorithm for Weighted K-means:GAWK) โดยมีจุดประสงค์เพื่อ 1) เทคนิค GA ค้นหาค่าถ่วงน้ำหนักคุณลักษณะข้อมูลที่เหมาะสมให้กับเทคนิคจัดกลุ่มเคมีน 2) เทคนิค GA ค้นหจำนวนกลุ่มที่เหมาะสมให้กับเทคนิคจัดกลุ่มเคมีน 3) การนำโมเดล GAWK ประยุกต์ใช้กับเว็บไซต์ให้บริการคำปรึกษาการเลือกสมัครงานแก่บัณฑิต

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ออกแบบโมเดล GAWK โดยการนำเทคนิค GA และเทคนิคจัดกลุ่มเคมีนมาประยุกต์ทำงานร่วมกัน ซึ่งอัลกอริทึมเทคนิคในโมเดล GAWK มีความแตกต่างจากอัลกอริทึมเทคนิคแบบดั้งเดิมบางส่วนผู้วิจัยขอกล่าวถึงรายละเอียดอัลกอริทึมเทคนิคเบื้องต้นก่อน

2.1 การทำเหมืองข้อมูล

การทำเหมืองข้อมูลเป็นกระบวนการค้นกรองข้อมูลจากฐานข้อมูลขนาดใหญ่โดยมีเทคนิคการจัดกลุ่มเคมีนเป็นเทคนิค

การทำเหมืองข้อมูลซึ่งมีความเหมาะสมในเรื่องการประมวลผลที่รวดเร็วและใช้ทรัพยากรในการคำนวณที่น้อย [1]

2.2 การจัดกลุ่มเคมีน

การจัดกลุ่มเคมีนมีวิธีแบ่งกลุ่มข้อมูลที่เหมือนกันอยู่ด้วยกัน โดยเป็นการแบ่งกลุ่มข้อมูลให้ได้จำนวนกลุ่ม k กลุ่ม โดยถูกกำหนดโดยผู้ใช้ และหาระยะทาง (Distance) จากสมการที่ (1)

$$d(x, y) = \sqrt{\sum_{i=1}^N (x_i - y_i)^2} \quad (1)$$

จากสมการที่ (1) กำหนดให้ $d(x, y)$ คือระยะห่างระหว่างตำแหน่ง (x_i, y_i) ใดๆ ในฐานข้อมูล N คือจำนวนข้อมูลเป็นเลขจำนวนเต็ม การทำงานอัลกอริทึมของการจัดกลุ่มเคมีนมีขั้นตอนดังนี้ 1) ทำการสุ่มข้อมูลออกมา k ตัว เพื่อเป็นศูนย์กลางกลุ่มเริ่มต้น 2) อ่านข้อมูลทั้งหมด จัดให้ข้อมูลแต่ละเรคคอร์ดอยู่ในกลุ่มให้ค่าระยะห่างกับเรคคอร์ดนั้นน้อยที่สุด 3) หากจุดศูนย์กลางกลุ่มใหม่โดยเฉลี่ยจากข้อมูลที่อยู่ในกลุ่มนั้น 4) ทำซ้ำในกระบวนการที่จนกระทั่งได้จุดศูนย์กลางของกลุ่ม 1 ข้อมูลเดิมและสมาชิกในกลุ่มไม่เปลี่ยน

2.3 ค่าน้ำหนัก

คุณลักษณะของข้อมูลแต่ละด้านมีอิทธิพลต่อผลลัพธ์ไม่เท่ากัน การกำหนดค่าน้ำหนักคุณลักษณะของข้อมูลแต่ละด้านก่อนกระบวนการจัดกลุ่มเคมีนจึงเป็นการเพิ่มประสิทธิภาพแก่ผลลัพธ์[3] และสามารถคำนวณได้จากสมการที่ (2)

$$dw(x, y) = \sqrt{\sum_{i=1}^N (wx_i - wy_i)^2} \quad (2)$$

จากสมการที่ (2) กำหนดให้ $dw(x, y)$ คือค่าน้ำหนักระยะห่างระหว่างตำแหน่ง (x_i, y_i) ใดๆ ในฐานข้อมูล, N คือจำนวนข้อมูลเป็นเลขจำนวนเต็ม, w_i คือ ค่าน้ำหนักที่กำหนดให้ i คือ จำนวนจริงมีค่าตั้งแต่ 0 ถึง 1

2.4 ขั้นตอนวิธีเชิงพันธุกรรม (Genetic Algorithm)

ขั้นตอนวิธีเชิงพันธุกรรมคืออัลกอริทึมเพื่อค้นหาการเพิ่มประสิทธิภาพที่ดีที่สุดด้วยการปรับตัว (Adaptive Global Optimization) โดยอัลกอริทึมมีขั้นตอนจำลองหลักการคัดเลือกตามธรรมชาติที่ต้องผ่านสภาพแวดล้อมธรรมชาติที่เหมาะสม (Fitness Function) ด้วยการวิวัฒนาการพันธุกรรม มีขั้นตอนต่างๆ เช่น การคัดเลือกพันธุกรรม (Selection) การสลับ

สายพันธุ (Crossover) การกลายพันธุ (Mutation) [4][7]

เทคนิคGAมีข้อดีในการค้นหาข้อมูลขนาดใหญ่และค้นหาโดเมนข้อมูลที่มีความอิสระระหว่างกันในแต่ละด้านของข้อมูล จึงถูกนำไปประยุกต์ใช้ร่วมกับการทำงานหลากหลายแขนงเช่น การจดจำรูปแบบ (Pattern Recognition) เครื่องมือเรียนรู้ (Machine Learning) การออกแบบวงจรVLSI [8]

2.5 การตรวจสอบการจัดกลุ่ม (Validation Clustering)

มาตรวัดที่ใช้บอกคุณภาพในการจัดกลุ่มว่าการจัดกลุ่มลักษณะใดดีกว่ากันสามารถวัดได้จากค่าผลรวมความผิดพลาด SSE (Sum of Squares Error) หากมีค่าน้อยจะแสดงถึงการจัดกลุ่มมีประสิทธิภาพที่ดี ดังสมการที่ (3) ซึ่งเป็นผลรวมของค่า WSS(Within cluster Sum of Squares) ดังสมการที่(4) เป็นผลรวมค่าระยะห่างระหว่างสมาชิกภายในกับศูนย์กลาง กลุ่ม และค่า BSS(Between cluster Sum of Squares) ดังสมการที่ (5) เป็นผลรวมระยะห่างระหว่างกลุ่ม

$$SSE = WSS + BSS \quad (3)$$

$$WSS = \sum_i \sum_{x \in C_i} (x - m_i)^2 \quad (4)$$

$$BSS = \sum_{x \in C_i} |C_i| (m - m_i)^2 \quad (5)$$

3. การหาค่าถ่วงน้ำหนักสำหรับเคมีนโดยใช้ขั้นตอนวิธีเชิงพันธุกรรม (GAWK)

งานวิจัยนี้ผู้วิจัยได้ประยุกต์นำเทคนิค GA และเทคนิคจัดกลุ่มเคมีนทำงานร่วมกัน เรียกโมเดลนี้ว่าการหาค่าถ่วงน้ำหนักสำหรับเคมีนโดยใช้ขั้นตอนวิธีเชิงพันธุกรรม โดยมีจุดประสงค์เพื่อ 1) เทคนิค GA สามารถค้นหาค่าถ่วงน้ำหนักคุณลักษณะข้อมูลที่เหมาะสมให้ให้กับเทคนิคจัดกลุ่มเคมีน 2) เทคนิค GA สามารถค้นหาจำนวนกลุ่มที่เหมาะสมให้กับเทคนิคจัดกลุ่มเคมีน 3) การนำโมเดล GAWK ถูกประยุกต์ใช้กับเว็บไซต์ให้บริการคำปรึกษาการเลือกสมัครงานแก่บัณฑิตจบใหม่

3.1 ขั้นตอนในการสร้างโมเดลการหาค่าถ่วงน้ำหนักสำหรับเคมีนโดยใช้ขั้นตอนวิธีเชิงพันธุกรรม

ผู้วิจัยได้ออกแบบอัลกอริทึมเพื่อแสดงขั้นตอนในการสร้างโมเดล GAWK ได้แสดงไว้ดังแผนผังการทำงานอัลกอริทึมในภาพที่1 ประกอบด้วยขั้นตอนดังนี้

- 1) รูปแบบโครโมโซม (Chromosome Encoding) คือ การ

กำหนดรูปแบบโครงสร้างโครโมโซมตามน้ำหนักคุณลักษณะของข้อมูลของบัณฑิตได้แก่ หลักสูตร เพศ ประเภทตำแหน่งงาน ประเภทหน่วยงาน จังหวัดที่ตั้งบริษัท เป็นคุณลักษณะที่เป็นปัจจัยส่งผลต่อเงินเดือนที่ได้รับดังตารางที่ 2 ซึ่งสามารถแบ่งคุณลักษณะตัวแปรนำเข้าที่สำคัญได้จำนวน 5 คุณลักษณะ ดังนั้น ผู้วิจัยได้เริ่มกำหนดรูปแบบโครงสร้างของโครโมโซมด้วยไบนารีสตริง (Binary String) ประกอบด้วย 5 บิตสตริง



ภาพที่ 1: แผนผังแสดงโมเดล GAWK

2) ประชากรต้นกำเนิด (Initial Population) คือ การกำหนดประชากรก่อนเข้ากระบวนการขั้นตอนวิธีเชิงพันธุกรรม ประชากรต้นกำเนิดจะได้รับการสุ่ม (Random) จากประชากรทั้งหมด โดยผู้วิจัยได้กำหนดประชากรเริ่มต้นจำนวน 32 โครโมโซม ตามเซตของโอกาสเกิดทั้งหมดตามคุณลักษณะที่ส่งผลจะมี 32 กรณี และในการสุ่มจะต้องสุ่มให้ได้จำนวนเท่ากับขนาดของรุ่นที่ได้กำหนดไว้

3) ค่าความเหมาะสม (Fitness Function) คือโครโมโซมทุกตัวต้องมีค่าซึ่งบอกถึงความเหมาะสมที่จะพิจารณาว่าสมควรสืบสายพันธุต่อหรือไม่ โดยใช้สมการหาค่าความเหมาะสมที่สอดคล้องกับปัญหา โดยจุดประสงค์ (Objective Function) ของการศึกษางานวิจัยครั้งนี้คือการค้นหาค่าถ่วงน้ำหนักคุณลักษณะข้อมูลและค้นหาจำนวนกลุ่มที่เหมาะสมให้กับเทคนิคจัดกลุ่มเคมีน โดยพิจารณาที่ทำให้ประสิทธิภาพความ

ถูกต้องมีค่ามากขึ้นวัดโดยใช้สมการที่ (6)

$$\max(dw(x, y)) = \max \sqrt{\sum_{i=1}^N w(x_i - y_i)^2} \quad (6)$$

จากสมการที่(6)กำหนดให้ $\max(dw(x, y))$ คือค่าค้นหาค่าความถูกต้องที่ดีที่สุดในการจัดกลุ่ม ที่ผ่านโมเดลการหาค่าถ่วงน้ำหนักสำหรับเคมีน โดยใช้ขั้นตอนวิธีเชิงพันธุกรรมที่ตำแหน่ง (x_i, y_i) ใดๆ ในฐานข้อมูล N ข้อมูลและถ่วงน้ำหนักด้วย w มีค่าระหว่าง 0 ถึง 1

4) การดำเนินการทางพันธุกรรม (Genetic Operation) คือกระบวนการที่สำคัญของ ขั้นตอนวิธีเชิงพันธุกรรม ประกอบด้วย a) ขั้นตอนการคัดเลือก (Selection) เป็นขั้นตอนคัดเลือกโครโมโซมเพื่อที่จะนำมาเป็นพ่อแม่ในการสืบสายพันธุ์ b) ขั้นตอนการสลับสายพันธุ์(Crossover) เป็นการสลับสายพันธุ์ระหว่าง 2 โครโมโซมให้ได้โครโมโซมใหม่ โดยผู้วิจัยได้เลือกวิธีการสุ่มตำแหน่งเพียงหนึ่งตำแหน่ง (Single Point) และทำการคัดลอกรหัสโครโมโซมทุกอย่างที่อยู่หน้าตำแหน่งสลับสายพันธุ์ของโครโมโซมพ่อ แล้วคัดลอกรหัสโครโมโซมทุกอย่างที่อยู่หลังตำแหน่งสลับสายพันธุ์ของโครโมโซมแม่ ซึ่งจะได้โครโมโซมลูกตัวที่หนึ่ง จากนั้นทำการคัดลอกรหัสโครโมโซมทุกอย่างที่อยู่หน้าตำแหน่งสลับสายพันธุ์ของโครโมโซมแม่ แล้วคัดลอกรหัสโครโมโซมทุกอย่างที่อยู่หลังตำแหน่งสลับสายพันธุ์ของโครโมโซมพ่อ ซึ่งจะได้โครโมโซมลูกตัวที่สอง และค่าความเหมาะสมของโครโมโซมที่เกิดใหม่ควรลักษณะโครโมโซมคล้ายพ่อแม่ โดยกำหนดค่าความเหมาะสมของโครโมโซมลูกอยู่ไม่เกิน 1.2 ของพ่อแม่ c) ขั้นตอนการกลายพันธุ์ (Mutation) เป็นการนำโครโมโซมลูกจากขั้นตอนก่อนหน้า ทำการสุ่มตำแหน่งเพื่อเปลี่ยนค่ายีนในโครโมโซม โดยเปลี่ยนจาก 1 เป็น 0 หรือ จาก 0 เป็น 1 และโอกาสของการกลายพันธุ์อยู่ที่ค่าความน่าจะเป็น 0.01

5) การกำหนดน้ำหนักและจำนวนกลุ่ม คือ หลังจาก GA ค้นหาถ่วงน้ำหนักคุณลักษณะข้อมูลที่เหมาะสม โดยกำหนดให้มีการแปลงน้ำหนักคุณลักษณะทั้ง 5 ด้านมีค่าตัวเลขทศนิยมและมีค่ารวมกันมีค่าเท่ากับ 1 โดยและค้นหาจำนวนกลุ่มที่เหมาะสมโดยผู้วิจัยได้กำหนดให้มีค่าเป็นเลขจำนวนเต็ม ซึ่งขั้นตอนนี้ทำให้ได้ค่าน้ำหนักและกลุ่มก่อนการใช้เทคนิคจัดกลุ่มเคมีนเป็นขั้นตอนถัดไป

6) การจัดกลุ่มเคมีน คือ ข้อมูลฝึกสอนจำนวนสองในสามส่วนของข้อมูลบัพทิตทั้งหมดซึ่งมีคุณลักษณะของบัพทิต ได้แก่ หลักสูตร เพศ ประเภทตำแหน่ง ประเภทหน่วยงาน จังหวัดที่ตั้งที่ทำงาน เป็นตัวแปรนำเข้าและระดับเงินเดือนที่ได้รับดังตารางที่ 1 เป็นผลลัพธ์ คุณลักษณะของข้อมูลนำเข้าทั้ง 5 คุณลักษณะจะถูกจะถ่วงน้ำหนักที่เหมาะสมด้วยเทคนิค GA ที่ได้ค้นหาให้ และจำนวนกลุ่มที่เหมาะสมในการแบ่งประเภทได้รับจากเทคนิค GA ค้นหาให้เช่นกันจากขั้นตอนก่อนหน้า

ตารางที่ 1: การกำหนดระดับเงินเดือนบัพทิต

กลุ่ม	รายได้ต่อเดือน	คน
1	รายได้ต่ำกว่า 7,500 บาท	141
2	รายได้ระหว่าง 7,500-10,000 บาท	231
3	รายได้ระหว่าง 10,000-12,500 บาท	195
4	รายได้ระหว่าง 12,500-15,000 บาท	135
5	รายได้มากกว่า 15,000 บาท	216

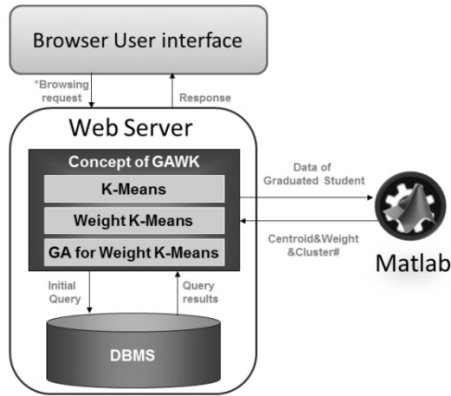
ซึ่งในขั้นตอนนี้จะได้ตัวแทนข้อมูลของทุกกลุ่ม (Centroid) เป็นตัวแทนข้อมูลสมาชิกที่ใช้ในการฝึกสอน และ เนื่องจากระดับเงินเดือนกับจำนวนกลุ่มที่แบ่งแยกประเภทมีค่าไม่เท่ากัน ผู้วิจัยจึงได้กำหนดให้ในการจัดกลุ่มใดๆ หากระดับเงินเดือนใดเป็นผลลัพธ์ของข้อมูลสมาชิกที่มากที่สุด ระดับเงินเดือนนั้นจะเป็นตัวแทนข้อมูลผลลัพธ์ในการจัดกลุ่มเคมีนกลุ่มนั้น

7) การวัดค่าความถูกต้อง คือ การนำข้อมูลทดสอบจำนวนหนึ่งในสามของข้อมูลบัพทิตทั้งหมดถูกถ่วงน้ำหนักที่เหมาะสมด้วยเทคนิค GA ที่ได้ค้นหาให้ หากข้อมูลนั้นใกล้กับตัวแทนข้อมูล (Centroid) ของกลุ่มใดผลลัพธ์ของข้อมูลจะได้ระดับเงินเดือนที่เป็นผลลัพธ์ในการจัดกลุ่มเคมีนกลุ่มนั้น และผู้วิจัยได้เปรียบเทียบข้อมูลระหว่าง ผลลัพธ์ระดับเงินเดือนของบัพทิตที่ได้จากทำโดยโมเดล GAWK และระดับเงินเดือนของบัพทิตที่ได้จากแบบสอบถามจริง เพื่อวัดประสิทธิภาพด้านความถูกต้อง (Accuracy) ของโมเดล

8) เงื่อนไขการหยุด (Stopping Criteria) คือการกำหนดเงื่อนไขให้หยุดเทคนิคGA โดยผู้วิจัยได้กำหนดเงื่อนไขการหยุดไว้เมื่อ เมื่อสร้างโครโมโซมลูกถึงรุ่น (Generation) ที่ 25 หรือค่าความเหมาะสมมีค่าค้างค่าเดิม (Stall Generation) เกิน 5 รุ่น

3.2 สถาปัตยกรรมโครงสร้างเว็บไซต์ GAWK

การนำโมเดล GAWK นำไปประยุกต์กับเว็บไซต์ให้บริการคำปรึกษา ได้แสดงสถาปัตยกรรมโครงสร้างเว็บไซต์เหมือนข้อมูลในภาพที่ 2 ประกอบด้วยโครงสร้างดังนี้



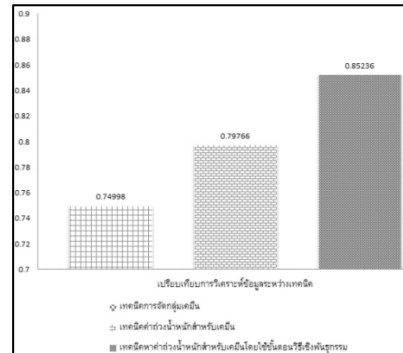
ภาพที่ 2: สถาปัตยกรรมโครงสร้างเว็บไซต์เหมือนข้อมูล

- 1) Browser User Interface คือส่วนติดต่อกับผู้ใช้เพื่อบริการคำปรึกษาการเลือกสมัครงานแก่บัณฑิตจบใหม่
- 2) Web Server คือ เว็บไซต์ให้บริการปรึกษาโดยมีระบบจัดการฐานข้อมูลDBMSเก็บและค้นหาข้อมูลสถิติของบัณฑิต
- 3) MATLAB Engine คือ ส่วนประมวลผลโดยนำข้อมูลสถิติบัณฑิตคำนวณด้วยโมเดล GAWK ผลลัพธ์ที่ได้จากองค์ประกอบส่วนนี้คือ ตัวแทนข้อมูลของทุกกลุ่ม (Centroid) นำให้นักคุณลักษณะข้อมูลที่เหมาะสม และ ระดับเงินเดือนที่เป็นตัวแทนของกลุ่มข้อมูล

4. ผลการวิจัย

ข้อมูลแบบสอบถามภาวะการมีงานทำของบัณฑิตจบใหม่ของมหาวิทยาลัยเทคโนโลยีราชมงคลจากด้านนา เขตพื้นที่ตาก 918 แบบสอบถาม นำเข้าสู่กระบวนการในการสกัดข้อมูลเพื่อจัดข้อมูลรบกวน (Noisy Data) จากนั้นนำข้อมูลที่ผ่านการสกัดแล้วมาแบ่งเป็นข้อมูลสำหรับสอน (Training Data) จำนวน 2/3 ส่วน และข้อมูลสำหรับทดสอบ (Test Data) จำนวน 1/3 ส่วน มาทำการทดลองโดยใช้อัลกอริทึมตามสมการที่ (1) (2) (6) เพื่อเป็น โมเดลในการทดลองเปรียบเทียบประสิทธิภาพในการการวิเคราะห์ข้อมูลเพื่อทำนายโดยใช้สมการ(3) (4) (5) ในระดับเงินเดือนบัณฑิตสำเร็จใหม่ระหว่างเทคนิคการจัดกลุ่มเคมีน เทคนิคค่าถ่วงน้ำหนักสำหรับ

เคมีน และเทคนิคค่าถ่วงน้ำหนักสำหรับเคมีน โดยใช้เทคนิค GA โดยการวิจัยครั้งนี้ผู้วิจัย ได้ใช้โปรแกรม Matlab 2012b เป็นซอฟต์แวร์ช่วยในการประมวลผลข้อมูล ซึ่งทำการทดลอง 10 ครั้ง มีผลการทดลองเปรียบเทียบเทคนิคทั้งสามอัลกอริทึมได้ค่าเฉลี่ยดังภาพที่ 3



ภาพที่ 3: เปรียบเทียบการวิเคราะห์ข้อมูล

ผลการทดลองด้วยเทคนิคการจัดกลุ่มเคมีน กับการจัดกลุ่มเงินเดือน ของกลุ่มระดับเงินเดือนต่างๆ ที่ได้ทำการสกัดข้อมูลไว้แล้ว มีค่า Accuracy มีค่าต่ำที่สุดอยู่ที่ประมาณ 74.9% มีค่า SSE อยู่ระดับที่ 172.1, WSS อยู่ที่ระดับที่ 378.1 และBSS อยู่ระดับที่ 550.3 ซึ่งอยู่ในระดับที่สูงที่สุด เนื่องจากข้อมูลที่ได้นำมาทดลองเป็นข้อมูลแต่ละด้านมีความสำคัญเท่ากัน นั่นจึงทำให้เทคนิคการจัดกลุ่มเคมีน มีค่า Accuracy อยู่ในระดับต่ำที่สุด และ SSE WSS BSS และอยู่ในระดับสูงที่สุด

ผลการทดลองการใช้เทคนิคค่าถ่วงน้ำหนักสำหรับเคมีน วิเคราะห์ข้อมูล มีประสิทธิภาพในการทำนายระดับเงินเดือนที่ ดีกว่าเทคนิคการจัดกลุ่มเคมีนค่า Accuracy อยู่ที่ 79.8% มีค่า SSE อยู่ระดับที่ 82.7, WSS อยู่ที่ระดับที่ 311.5 และBSS อยู่ระดับที่ 394.1 กล่าวได้ว่าการกำหนดค่าถ่วงน้ำหนักที่เหมาะสม มีผลต่อการเพิ่มประสิทธิภาพต่อเทคนิคการจัดกลุ่มเคมีน

สำหรับการทดลองการใช้เทคนิค GAWK วิเคราะห์ข้อมูล ได้ค่าจำนวนจัดกลุ่มที่เหมาะสมเป็น 4 และ ค่าน้ำหนักของลักษณะทั้งห้าด้านมีค่าดังนี้ ค่าน้ำหนักหลักสูตร 0.2466, ค่าน้ำหนักเพศ 0.1246, ค่าน้ำหนักประเภทตำแหน่งงาน 0.2262, ค่าน้ำหนักประเภทหน่วยงาน 0.1561, ค่าน้ำหนักจัดหมวดหมู่ที่ตั้งบริษัท 0.2464 เมื่อทำการทดลองทำให้ค่า Accuracy มากที่สุดเป็น 85.24% และเพิ่มมากกว่าค่า Accuracy ที่ได้จากเทคนิคการจัดกลุ่มเคมีนแบบเดิม 10% โดยมีค่า SSE อยู่ระดับที่ 54.6, WSS อยู่ที่ระดับที่ 276.6 และBSS อยู่ระดับที่ 331.27 ซึ่ง

มีค่าต่ำที่สุดแสดงถึงการจัดกลุ่มมีประสิทธิภาพดีที่สุดจากผลการทดลองดังกล่าวพบว่าเทคนิค GAWK เมื่อวัดประสิทธิภาพเพื่อทำนายในระดับเงินเดือนบัณฑิตสำเร็จใหม่ให้ค่าความถูกต้องแม่นยำมากกว่า โมเดลในการทดลองเปรียบเทียบกับประเภทอื่น ระหว่างเทคนิคการจัดกลุ่มเคมีน และเทคนิคค่าถ่วงน้ำหนักสำหรับเคมีน

5. สรุปผล และข้อเสนอแนะ

งานวิจัยนี้มีวัตถุประสงค์นำเสนอวิธีการจัดกลุ่มข้อมูลโดยจัดกลุ่มข้อมูลได้นำเทคนิคจัดกลุ่มเคมีน มีกำหนดค่าน้ำหนักคุณลักษณะที่เหมาะสมและจำนวนกลุ่มที่เหมาะสมด้วยขั้นตอนวิธีเชิงพันธุกรรม เป็นการเพิ่มประสิทธิภาพด้านความถูกต้องให้สูงขึ้น ผู้วิจัยได้ออกแบบโมเดลการหาค่าถ่วงน้ำหนักสำหรับเคมีน โดยใช้ขั้นตอนวิธีเชิงพันธุกรรม GAWK

จากการวิจัยการจัดกลุ่มข้อมูลภาวะการมีงานทำของบัณฑิตโดยโมเดล GAWK พบว่าการแบ่งกลุ่มตามระดับเงินเดือนด้วยเทคนิคจัดกลุ่มเคมีนมีค่าความถูกต้อง 74.99% เทคนิคค่าถ่วงน้ำหนักสำหรับเคมีนมีค่าความถูกต้อง 79.77% และหลังจากใช้โมเดล GAWK มีค่าความถูกต้อง 85.24 % แสดงให้เห็นว่าถ้ามีการกำหนดค่าน้ำหนักที่เหมาะสมให้กับการจัดกลุ่มเคมีน ทำให้การจัดกลุ่มข้อมูลมีความถูกต้องมากยิ่งขึ้น

ผู้วิจัยสรุปได้ว่าเทคนิค GAWK ได้ผลลัพธ์ประสิทธิภาพด้านความถูกต้องดีขึ้นเมื่อเปรียบเทียบกับการใช้เทคนิคจัดกลุ่มเคมีน เพราะการถ่วงน้ำหนักสำคัญให้คุณลักษณะข้อมูลมีผลต่อทำนายผลที่แม่นยำขึ้นซึ่งผลลัพธ์จากการวิจัยมีความสอดคล้องในแนวทางเดียวกันกับงานวิจัย [4][5] และ เทคนิคGAWK จะค้นหาน้ำหนักความสำคัญคุณลักษณะของข้อมูลที่เหมาะสมให้กับเทคนิคจัดกลุ่มเคมีนซึ่งได้ผลลัพธ์ประสิทธิภาพด้านความถูกต้องดีขึ้น ซึ่งผลลัพธ์จากการวิจัยมีความสอดคล้องในแนวทางเดียวกันกับงานวิจัย[3]

สถาปัตยกรรมโครงสร้างเว็บไซต์ GAWK ถูกนำไปประยุกต์ใช้กับการทำเหมืองข้อมูลบนเว็บไซต์ที่มีข้อมูลในการวิเคราะห์ในปริมาณที่มาก เพื่อสกัดข้อมูลที่มีลักษณะคล้ายกันจัดกลุ่มเดียวกันซึ่งจะได้ตัวแทนข้อมูลที่แทนข้อมูลของสมาชิกทั้งหมด ทำให้ผลลัพธ์ในการค้นหาข้อมูลบนเว็บไซต์มีความรวดเร็ว ขณะนี้ผู้วิจัยได้ออกแบบการทำงานเว็บไซต์โดยใช้สถาปัตยกรรมโครงสร้างเว็บไซต์โดยโมเดล GAWK ดังภาพที่

4 เพื่อให้บริการคำปรึกษาช่วยในการตัดสินใจในการเลือกสมัครงานในตำแหน่งที่เหมาะสมแก่บัณฑิตจบใหม่ของมหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เขตพื้นที่ตาก ประเด็นในการศึกษาวิจัยต่อการปรับปรุงโมเดลที่สามารถเรียนรู้ข้อมูลใหม่เองทันทีได้โดยปราศจากผู้เชี่ยวชาญในการปรับปรุงองค์ความรู้ และการจัดการตัวแทนข้อมูลหลังจากเพิ่มข้อมูลองค์ความรู้ใหม่

ภาพที่ 4: เว็บไซต์ให้คำปรึกษาการสมัครงานโดยโมเดล GAWK

เอกสารอ้างอิง

- [1] O. H. Odukoya, G. A. Aderounmu, "An Improved Data Clustering Algorithm for Mining Web Documents," in *2010 International Conference on Computational Intelligence and Software Engineering*, 2010, pp. 1–8.
- [2] A. Kumar, R. Sinha, V. Bhattacharjee, D. S. Verma, "Modeling using K-means clustering algorithm," in *2012 1st International Conference on Recent Advances in Information Technology (RAIT)*, 2012, pp. 554–558.
- [3] R. C. de Amorim, "Constrained clustering with Minkowski Weighted K-Means," in *2012 IEEE 13th International Symposium on Computational Intelligence and Informatics (CINTI)*, 2012, pp. 13–17.
- [4] W. Min and Y. Siqing, "Improved K-means clustering based on genetic algorithm," in *2010 International Conference on Computer Application and System Modeling (ICCSM)*, 2010, vol. 6, pp. V6–636–V6–639.
- [5] F.-X. Wu, "A genetic weighted k-means algorithm for clustering gene expression data," in *Second International Multi-Symposiums on Computer and Computational Sciences*, 2007, pp. 68–75.
- [6] S. Arunprasath, S. Chandrasekar, K. Venkatalakshmi, and S. M. Shalinie, "Classification of remote sensed image using Rapid Genetic k-Means algorithm," in *2010 IEEE International Conference on Communication Control and Computing Technologies (ICCCCT)*, 2010, pp. 677–682.
- [7] M. Laszlo, "A genetic algorithm that exchanges neighboring centers for k-means clustering," *Pattern Recognit. Lett.*, vol. 28, no. 16, pp. 2359–2366, Dec. 2007.

- [8] K. Krishna and M. N. Murty, "Genetic K-means algorithm," *IEEE Trans. Syst. Man Cybern. Part B Cybern.*, vol. 29, no. 3, pp. 433–439, 1999.

ระบบการจำแนกพันธุ์ยางพาราโดยใช้ตัวจำแนกนาอิวเบย์

Classification System Para Rubber Varieties using Naïve Bayes

วรรณสิริ ฐระชน (Wannasiri Thurachon)¹ วรพจน์ สุเมธาวัฒนพงศ์ (Worraphot Sumethawatthanaphong)²

และ นัทธวิภา ส่องสุข (Nutthavipa Songsook)³

สาขาวิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏอุดรธานี
kaitock@hotmail.com, worraphot.dbz@gmail.com and nutthavipa_dbz@hotmail.com

บทคัดย่อ

การปลูกยางพาราที่ไม่เหมาะสมกับลักษณะประจำพันธุ์ของยางพาราในพื้นที่เพาะปลูกของเกษตรกร เป็นปัญหาหนึ่งในการทำสวนยางพาราของเกษตรกร ก่อให้เกิดความยุ่งยากในการบำรุงรักษา รวมทั้งคุณภาพและปริมาณผลผลิตที่ได้ไม่คุ้มค่ากับการลงทุน สาเหตุหลักเกิดจากเกษตรกรไม่มีความรู้ในการคัดเลือกพันธุ์ยางพาราที่เหมาะสมมาปลูกในพื้นที่ของตน และการได้ต้นยางพาราที่ผิดพันธุ์ไปปลูก วัตถุประสงค์ของการวิจัยในครั้งนี้ เพื่อสร้างระบบการจำแนกพันธุ์ยางพาราโดยใช้ตัวจำแนกนาอิวเบย์ โดยใช้ข้อมูลตัวอย่างพันธุ์ยางพาราที่มีการปลูกมากในภาคตะวันออกเฉียงเหนือของประเทศไทย จำนวน 5 พันธุ์ โดยอาศัยคุณลักษณะเด่นประจำพันธุ์ของยางพารา 6 คุณลักษณะ คือ ฉัตร ใบ ก้านใบ ก้านใบย่อย เปลือก และตา ซึ่งมีคุณลักษณะย่อยอื่นๆ ที่ใช้ในการจำแนกพันธุ์ยางพารา โดยระบบสามารถทำการจำแนกพันธุ์ยางพาราที่ได้ผลลัพธ์ และระดับความถูกต้องของการจำแนกพันธุ์ยางพาราเท่ากับ 98.02%

คำสำคัญ: การจำแนกพันธุ์ยางพารา, การคัดเลือกพันธุ์ยางพารา, นาอิวเบย์, ยางพารา

Abstract

Para rubber planting that is inappropriate with features of Para rubbers varieties in planting areas of agriculturists is one of the problems in Para rubber gardening. This causes a difficulty in maintenance and the received quality and productivity are not worthwhile for investment. The main reason are agriculturists have no skills in selecting proper Para rubber varieties to be planted in their areas and bringing wrong Para rubber varieties to be planted. The purpose of this research is to

create a classification system of Para rubber varieties by using the Naive Bayes technique based on sample data of Para rubber varieties which are planted abundantly in the Northeast of Thailand. This sample data comprises 5 varieties of Para rubber based on 6 features of Para rubber varieties, that is, Leaf storey, Leaf, Petiole, Petiolules, Bark and Axillary bud, which have sub-features for using in the classification of Para rubber varieties. The classification system can classify Para rubber varieties and indicate results. The accuracy level of classification of Para rubber varieties equals to 98%.

Keywords: Classification of Para Rubber Varieties, Selection of Para Rubber, Naïve Bayes, Para rubber

1. บทนำ

ยางพาราเป็นหนึ่งในพืชเศรษฐกิจหลัก ที่มีความสำคัญต่อเศรษฐกิจไทยมานานนับร้อยปี โดยในปีที่ผ่านมายางพาราเป็นสินค้าส่งออกหลักของประเทศ มีพื้นที่ปลูกทั่วประเทศกว่า 14 ล้านไร่ ประชากรกว่า 6 ล้านคน อาศัยยางพาราเป็นหลักในการประกอบอาชีพ ทำให้ประเทศไทยเป็นผู้ผลิตยางพาราในสัดส่วนประมาณ 1 ใน 3 ของผลผลิตยางโลก และส่งออกยางพาราเป็นอันดับหนึ่งของโลก รัฐบาลมีนโยบายในการขยายเขตพื้นที่เพาะปลูกยางพาราเพิ่มขึ้นในพื้นที่ภาคเหนือและภาคตะวันออกเฉียงเหนือ ซึ่งรัฐบาลได้เข้ามาช่วยในทุกด้านของการทำสวนยางพารา รวมไปถึงการแจกต้นกล้าและกิ่งพันธุ์ของยางพาราให้แก่เกษตรกร ซึ่งบางครั้งพันธุ์ยางพาราที่ได้อาจไม่เหมาะสมกับพื้นที่เพาะปลูกของเกษตรกร ซึ่งสาเหตุเกิดมาจากการคัดเลือกพันธุ์ที่ไม่เหมาะสม การเพาะชำผิดพลาด บางพันธุ์อาจมีการเปลี่ยนแปลงไปตามสภาพแวดล้อมที่แตกต่างกัน หรือแม้กระทั่งลักษณะภูมิประเทศพื้นที่เพาะปลูกของเกษตรกรเอง

ปัญหาดังกล่าวส่งผลกระทบต่อภัยให้กับเกษตรกรชาวสวนยาง ก่อให้เกิดความยุ่งยากในการดูแลรักษา เนื่องจากยางพาราแต่ละพันธุ์จะตอบสนองต่อระบบการกรีด การต้านทานโรค และข้อจำกัดของสภาพพื้นที่แตกต่างกัน เมื่อต้นยางพาราโตเต็มที่พร้อมกรีดอาจทำให้ได้น้ำยางน้อยคุณภาพต่ำ ทำให้ไม่คุ้มค่าต่อการลงทุน ดังนั้นเพื่อให้เกษตรกรเลือกพันธุ์ยางพาราที่เหมาะสมมาปลูก จึงควรมีวิธีการจำแนกพันธุ์ยางพาราที่ถูกต้องเหมาะสม

วิธีการในการจำแนกพันธุ์ยางพารา ตามหลักทางการเกษตรนั้น สามารถทำได้ 2 วิธี [1] คือ การจำแนกพันธุ์ยางพาราด้วยสายตาโดยดูจากลักษณะเฉพาะประจำพันธุ์ของยางพารา เช่น ลักษณะของลำต้น กิ่ง ใบ ดอก รูปทรงฉัตร และวิธีที่สองคือการจำแนกพันธุ์ยางพาราทางวิทยาศาสตร์ โดยการจัดทำ ลายพิมพ์ DNA วิธีนี้มีความแม่นยำสูง แต่มีขั้นตอนการปฏิบัติที่ยุ่งยากใช้เวลาประมวลผลนาน และเสียค่าใช้จ่ายมาก ดังนั้นส่วนใหญ่นิยมใช้วิธีการจำแนกพันธุ์ยางพาราด้วยสายตามากกว่า เพราะเป็นวิธีที่ตรวจสอบได้ง่าย สะดวกรวดเร็ว สามารถบอกชนิดของพันธุ์ยางได้ทันที อย่างไรก็ตามวิธีวิธีการจำแนกโดยใช้สายตานั้นจำเป็นต้องใช้ผู้เชี่ยวชาญที่มีความรู้เกี่ยวกับลักษณะประจำพันธุ์ของยางพาราแต่ละพันธุ์มาช่วยในการจำแนกพันธุ์ยางพารา

ดังนั้น หากนำวิธีการจำแนกพันธุ์ยางพาราโดยใช้หลักการคำนวณทางคอมพิวเตอร์ด้วยวิธีการจำแนกข้อมูล (Data Classification) จะทำให้การจำแนกพันธุ์ยางพาราทำได้รวดเร็วและแม่นยำยิ่งขึ้น ซึ่งงานวิจัยในครั้งนี้อาศัยอัลกอริทึมนาอิวเบย์ (Naïve Bayes)[2], [3], [4] มาใช้ในการจำแนกพันธุ์ให้กับต้นยางพารา และผลที่ได้จะทำให้เกษตรกรมีความมั่นใจในการเลือก และจำแนกพันธุ์ยางพาราที่เหมาะสมมาปลูกในพื้นที่ของตนมากขึ้น

2. ทฤษฎีที่เกี่ยวข้อง

2.1 อัลกอริทึม นาอิวเบย์ (Naïve Bayes)

นาอิวเบย์เป็นเครื่องจักรเรียนรู้ที่อาศัยหลักการความน่าจะเป็น (probability) [2], [3], [4] ตามทฤษฎีของเบย์ (Bayes's theorem) ซึ่งมีอัลกอริทึมที่ไม่ซับซ้อน เป็นขั้นตอนวิธีในการจำแนกข้อมูล โดยการเรียนรู้ปัญหาที่เกิดขึ้น เพื่อนำมาสร้างเงื่อนไขการจำแนกข้อมูลใหม่ เป็นการจำแนกข้อมูลโดยใช้ความน่าจะเป็น และคำนวณการแจกแจงความน่าจะเป็นตามสมมติฐานที่ตั้งให้กับข้อมูล จากการคำนวณตัวอย่างใหม่ที่ได้

จะถูกนำมาปรับเปลี่ยนการแจกแจง ซึ่งมีผลต่อการเพิ่มหรือลดความน่าจะเป็นของข้อมูล ข้อมูลใหม่ที่เกิดขึ้นและตัวแบบที่ตั้งไว้ให้กับข้อมูล จะถูกปรับเปลี่ยนไปตามข้อมูลใหม่ โดยผนวกกับข้อมูลเดิมที่มี

หลักการของนาอิวเบย์ ใช้การคำนวณหาความน่าจะเป็นซึ่งถูกใช้ในการทำนายผล เป็นเทคนิคในการแก้ปัญหาแบบการจำแนกประเภทที่สามารถคาดการณ์ผลลัพธ์ได้ มันจะทำการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรเพื่อใช้ในการสร้างเงื่อนไขความน่าจะเป็นสำหรับแต่ละความสัมพันธ์ นาอิวเบย์เป็นวิธีจำแนกประเภทข้อมูลที่มีประสิทธิภาพ มีอัลกอริทึมในการทำงานที่ไม่ซับซ้อน เหมาะกับกรณีของเซตตัวอย่างที่มีจำนวนมาก และคุณสมบัติ(Attribute) ของตัวอย่างไม่ขึ้นต่อกันโดยกำหนดให้ความน่าจะเป็นของข้อมูลเท่ากับสมการที่ 1 [5], [6]

$$P(a_1, a_2, \dots, a_n | v_j) = P(v_j) \prod_{i=1}^n P(a_i | v_j) \quad (1)$$

โดยที่ $P(v_j)$ คือ ค่าความน่าจะเป็นของข้อมูลที่ให้คลาส

$P(a_i | v_j)$ คือ ค่าความน่าจะเป็นของข้อมูลคุณลักษณะ

ที่ i มีค่า a_i และให้คลาส v_j

$P(a_1, a_2, \dots, a_n | v_j)$ คือ ค่าความน่าจะเป็นของข้อมูลทดสอบที่มีค่าคุณลักษณะ a_1, a_2, \dots, a_n ที่จะให้คลาส v_j

กลุ่ม v_j สำหรับข้อมูลที่มีคุณสมบัติ n ตัว $X = \{ a_1, a_2, \dots, a_n \}$ หรือ ใช้สัญลักษณ์ว่า $P(a_1, a_2, \dots, a_n | v_j)$

โดยที่ \prod หมายถึง ผลคูณของค่า $(a_i | v_j)$ ทั้งหมด $i = 1, 2, 3, \dots, n$ และ $j = 1, 2, 3, \dots, n$

กำหนดให้ E แทนข้อมูลที่นำมาใช้ในการคำนวณการแจกแจงความน่าจะเป็น *posteriori probability* ของสมมติฐาน h คือ $Pr(h/E)$ ตามทฤษฎีของนาอิวเบย์ ดังสมการที่ 2 [3],[4],[7]

$$Pr(h/E) = \frac{Pr(E | h)Pr(h)}{Pr(E)} \quad (2)$$

$Pr(h)$ คือ ความน่าจะเป็นก่อนหน้าของสมมติฐาน h

$Pr(E)$ คือ ความน่าจะเป็นก่อนหน้าของชุดข้อมูลตัวอย่าง E

$Pr(h/E)$ คือ ความน่าจะเป็นของ h เมื่อรู้ E

$Pr(E/h)$ คือ ความน่าจะเป็นของ E เมื่อรู้ h

2.2 เทคนิคเลือกคุณลักษณะเด่น (Feature Selection)

จากคุณลักษณะย่อยของขงพยากรณ์จำนวน 34 คุณลักษณะ ที่เก็บรวบรวมมา ข้อมูลจะมีหลายประเภทแตกต่างกันไป ซึ่งบางคุณลักษณะอาจมีความสัมพันธ์กับคุณลักษณะอื่น อันจะมีผลทำให้การจำแนกเกิดความผิดพลาด ดังนั้นจึงต้องเลือกคุณลักษณะที่เป็นคุณลักษณะเด่นเท่านั้น ซึ่งในงานวิจัยนี้ใช้

1) วิธีสัมประสิทธิ์สหสัมพันธ์แบบสเปียร์แมน (Spearman's Correlation Coefficient) [8] หากความสัมพันธ์ระหว่างข้อมูลประเภท Ordinal

2) วิธีไคสแควร์ (Chi-Square) [9] หากความสัมพันธ์ระหว่างข้อมูลประเภท Nominal จะได้ค่าสำคัญการทดสอบ (Sig : Significance test) หากค่า Sig. > 0.9 ให้เลือกคุณลักษณะนั้น เพราะมีความสัมพันธ์กันในระดับสูง

2.2.1 สัมประสิทธิ์สหสัมพันธ์แบบสเปียร์แมน (Spearman correlation coefficient)

สัมประสิทธิ์สหสัมพันธ์แบบสเปียร์แมน [8] เป็นวิธีที่ใช้วัดความสัมพันธ์ระหว่างตัวแปรหรือข้อมูล 2 ชุด โดยที่ตัวแปรหรือข้อมูล 2 ชุด นั้นจะต้องอยู่ในรูปของข้อมูลในมาตราจัดอันดับ (Ordinal value) เช่น ระดับความพึงพอใจ (มากที่สุด/มาก/ปานกลาง/น้อย/น้อยที่สุด) ซึ่งมีจำนวนสถานะ 5 ค่า หรือ ขนาด (ใหญ่/กลาง/เล็ก) มีสถานะ 3 ค่า

ค่าสัมประสิทธิ์สหสัมพันธ์แบบสเปียร์แมน คำนวณจากสูตรดังสมการที่ (3)

$$r_s = 1 - \frac{6 \sum D^2}{N(N^2 - 1)} \quad (3)$$

เมื่อ r_s เป็นค่าสัมประสิทธิ์สหสัมพันธ์แบบสเปียร์แมน $\sum D^2$ เป็นผลรวมของกำลังสองของผลต่างของอันดับคะแนนแต่ละคู่

N เป็นขนาดของกลุ่มตัวอย่าง

2.2.2 การทดสอบไคสแควร์ (Chi-square test)

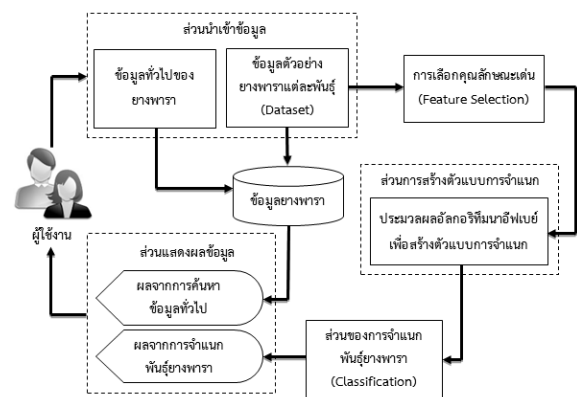
การทดสอบไคสแควร์ [9] สามารถทดสอบกับกลุ่มตัวอย่างได้หลายกลุ่ม และเป็นรูปแบบ nonparametric คือ ไม่คำนึงถึงการแจกแจงของประชากร ดังนั้นการทดสอบด้วยวิธีนี้จึงทำได้กว้างขวาง นอกจากนั้นเงื่อนไขในเรื่องของระดับการวัดนั้นการทดสอบไคสแควร์สามารถทำได้ทุกระดับ ในการทดสอบไคสแควร์จะใช้ข้อมูลแบบไบนารี (Binary value) เช่น จริง/เท็จ

หรือ ใช่/ไม่ใช่ หรือแบบนามมาตรา (Nominal value) เช่น สถานะ (โสด/แต่งงาน/หย่าร้าง) การคำนวณระยะห่างจะเป็นดังสมการที่ 4

$$K^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} \quad (4)$$

3. วิธีการดำเนินการวิจัย

การจำแนกพันธุ์ขงพยากรณ์ให้กับต้นขงที่ไม่ทราบพันธุ์ ก่อนอื่นต้องนำข้อมูลที่ได้จากแบบบันทึกคุณลักษณะของขงพยากรณ์ที่ได้เก็บมาจาก สถาบันวิจัยขงพยากรณ์ภาคตะวันออกเฉียงเหนือ นำมาแปลงให้อยู่ในรูปของตัวเลข จากนั้นนำไปผ่านกระบวนการเลือกคุณลักษณะ และสุดท้ายจะเป็นการสร้างตัวแบบการจำแนกพันธุ์ขงพยากรณ์โดยใช้ตัวจำแนกออนไลน์ แล้วนำตัวแบบการจำแนกพันธุ์นี้ไปจำแนกพันธุ์ขงพยากรณ์ที่ยังไม่ทราบพันธุ์ ซึ่งขั้นตอนได้แสดงไว้ดังภาพที่ 1



ภาพที่ 1 ภาพรวมระบบ (System Overview)

3.1 การรวบรวมข้อมูลพันธุ์ขงพยากรณ์

ในงานวิจัยนี้ได้ใช้ข้อมูลตัวอย่างพันธุ์ขงพยากรณ์ของสถาบันวิจัยขงพยากรณ์ภาคตะวันออกเฉียงเหนือ จังหวัดหนองคาย โดยจัดทำเป็นแบบบันทึกข้อมูลคุณลักษณะของขงพยากรณ์ มีผู้เชี่ยวชาญประจำสถาบันวิจัย ที่ทำงานด้านการจำแนกพันธุ์ขงพยากรณ์ที่มีประสบการณ์การทำงานมากกว่า 10 ปี จำนวน 3 ท่าน เป็นผู้เก็บข้อมูลคุณลักษณะของขงพยากรณ์จำนวน 5 พันธุ์ ละ 30 ต้น รวมเป็น 150 ต้น ที่เกิดจากการติดตามจากต้นขงแม่พันธุ์ และเป็นพันธุ์ขงที่มีนิยมปลูกกันมาก ในภาคตะวันออกเฉียงเหนือ ได้แก่ (1) พันธุ์ RRIT 251 (2) พันธุ์ RRIT 226 (3) พันธุ์ BPM 24 (4) พันธุ์ RRIM 600 และ (5) พันธุ์ PB 235 [10], [11] โดย

ข้อมูลคุณลักษณะของยางพาราที่เก็บนั้น มีการแยกพันธุ์ไว้ชัดเจนประกอบด้วย 6 คุณลักษณะหลักที่มีคุณลักษณะย่อย ดังแสดงในตารางที่ 1 และข้อมูลที่เกี่ยวข้องรวมมาจะไม่มีค่าว่างดังตัวอย่างข้อมูลที่แสดงในตารางที่ 2

ตารางที่ 1: คุณลักษณะประจำพันธุ์ยางพารา

คุณลักษณะหลัก	คุณลักษณะย่อย
(1) ฉัตร	ทรงฉัตร, ความกว้างฉัตร, ระยะห่างระหว่างฉัตร, การตกของใบในฉัตร
(2) ใบ	ขอบใบ, รูปร่างใบกลาง, สีใบ, ความเป็นมัน, ฐานใบ, ปลายใบ, สีของเส้นใบ, ผิวใบ, ใบกลางตัดตามขวาง, ใบกลางตัดตามยาว, ใบย่อยเปรียบเทียบกับใบกลาง, ตำแหน่งขอบใบย่อย, ความรู้สึกเมื่อสัมผัส, เส้นกลางใบ, ระดับของใบย่อย
(3) ก้านใบ	รูปร่างก้านใบ, ความยาวก้านใบ, รูปร่างฐานก้านใบ, ลักษณะก้านใบ, ทิศทางก้านใบทำกับลำต้น, รอยแผลก้านใบ
(4) ก้านใบย่อย	การแผ่ของก้านใบย่อย, การทำมุมระหว่างก้านใบย่อย, ความยาวก้านใบย่อย
(5) เปลือก	เปลือกส่วนสีเขียว, เปลือกส่วนสีน้ำตาล
(6) ตา	ทิศทางของตาคิ้ว, ลักษณะของตาก้านใบ, ที่ตั้งของตาก้านใบ, ลักษณะของตาคิ้ว

ตารางที่ 2: ตัวอย่างข้อมูลของยางพาราแต่ละพันธุ์

ต้นที่	1) ทรงฉัตร	2) ขอบใบ	...	34) ตาคิ้ว	ชื่อพันธุ์
1	รูปป้อม	เรียบ	...	เสมอลำต้น	RRIM 600
2	กรวย	เรียบ	...	เสมอลำต้น	RRIM 600
...
150	รูปป้อม	เรียบ	...	เสมอลำต้น	PB 235

3.2 การเลือกคุณลักษณะข้อมูล (Feature Selection)

จากคุณลักษณะของยางพาราที่เก็บรวบรวมมานั้น บางคุณลักษณะอาจมีความเกี่ยวข้องกับคุณลักษณะอื่น อันจะมีผลทำให้การจำแนกพันธุ์เกิดความผิดพลาด ดังนั้นจึงต้องเลือก

คุณลักษณะที่มีความเกี่ยวพันกันน้อยที่สุดออก ซึ่งมีหลายวิธีทางสถิติที่ใช้หาความเกี่ยวพันระหว่าง feature แต่ละคู่ ในงานวิจัยนี้ใช้วิธี Spearman's Correlation Coefficient [8] กับข้อมูลประเภท Ordinal และใช้วิธี Chi-Square [9] ในการหาความเกี่ยวพันระหว่างข้อมูลประเภท Nominal คุณลักษณะคู่ที่มีค่าสัมประสิทธิ์ความเกี่ยวพันกันมากกว่า 0.9 ตัวแทนของคู่ นั้นจะถูกเลือกเพื่อนำไปสร้างตัวแบบการจำแนกต่อไป ซึ่งจากการใช้ทั้งสองวิธี ทำให้ได้คุณลักษณะเด่นทั้งหมด 12 คุณลักษณะเด่น จากคุณลักษณะย่อยทั้งหมด 34 คุณลักษณะคือ ความกว้างของฉัตร, ความเป็นมัน, ฐานใบ, ผิวใบ, ใบตัดตามขวาง, ขอบใบ, ความยาวของก้านใบ, รูปร่างของฐานก้านใบ, รอยแผลก้านใบ, การทำมุมระหว่างก้านใบย่อย, ความยาวก้านใบย่อย และเปลือกส่วนสีเขียว จากการเลือกคุณลักษณะเด่นที่ได้จากสองวิธีข้างต้น ทำให้ได้คุณลักษณะเด่นทั้งหมด 12 คุณลักษณะ แต่การจำแนกพันธุ์ยางพาราในงานวิจัยนี้ เลือกใช้คุณลักษณะเด่นที่ให้ค่าความถูกต้องสูงสุด โดยได้ทำการทดลองในโปรแกรม Weka [12] ผลที่ได้คือ ค่าความถูกต้องสูงสุดของคุณลักษณะเด่น เท่ากับ 98% โดยใช้ลักษณะเด่นทั้งหมด 10 คุณลักษณะ ประกอบไปด้วย ความกว้างของฉัตร, ความเป็นมัน, ผิวใบ, ใบตัดตามขวาง, ขอบใบ, ความยาวของก้านใบ, รูปร่างของฐานก้านใบ, รอยแผลก้านใบ, ความยาวก้านใบย่อย และเปลือกส่วนสีเขียว โดยคุณลักษณะเด่นเหล่านี้จะนำไปใช้ในการสร้างตัวแบบการจำแนกพันธุ์ยางพาราต่อไป

3.3 การสร้างตัวแบบการจำแนกพันธุ์ยางพารา

ในการสร้างตัวแบบการจำแนกพันธุ์ยางพารานั้น ขั้นตอนแรกต้องทำการนับจำนวน dataset ของยางพาราแต่ละพันธุ์ว่ามีจำนวนเท่าไร โดยจำนวน dataset ทั้งหมดที่นำมาใช้ทำตัวแบบมีทั้งหมด 150 ต้น โดยจะทำการนับตามคุณลักษณะเด่นที่ได้เลือกทั้งหมด 10 คุณลักษณะ วิธีการนับคือ [3], [4], [7] ให้นับจำนวนของลักษณะย่อยของแต่ละพันธุ์ที่เกิดจากคุณลักษณะเด่นทั้ง 10 คุณลักษณะดังภาพที่ 2

ลักษณะใบตัดตามขวาง	ลักษณะขอบใบ	ลักษณะรอยแผลก้านใบ	ลักษณะการท่ามพรหวงก้านใบ	ลักษณะฐานใบ	====พันธุ์====
พันธุ์ RRIT251 = 23	พันธุ์ RRIT251 = 10	พันธุ์ RRIT251 = 8	พันธุ์ RRIT251 = 8	พันธุ์ RRIT251 = 25	พันธุ์ RRIT251 = 25
พันธุ์ RRIT226 = 7	พันธุ์ RRIT226 = 19	พันธุ์ RRIT226 = 4	พันธุ์ RRIT226 = 4	พันธุ์ RRIT226 = 19	พันธุ์ RRIT226 = 19
พันธุ์ BPM24 = 24	พันธุ์ BPM24 = 3	พันธุ์ BPM24 = 18	พันธุ์ BPM24 = 18	พันธุ์ BPM24 = 9	พันธุ์ BPM24 = 9
พันธุ์ RRM600 = 1	พันธุ์ RRM600 = 31	พันธุ์ RRM600 = 5	พันธุ์ RRM600 = 5	พันธุ์ RRM600 = 3	พันธุ์ RRM600 = 3
พันธุ์ PB235 = 7	พันธุ์ PB235 = 31	พันธุ์ PB235 = 24	พันธุ์ PB235 = 24	พันธุ์ PB235 = 24	พันธุ์ PB235 = 24
พันธุ์ RRIT251 = 6	พันธุ์ RRIT251 = 19	พันธุ์ RRIT251 = 1	พันธุ์ RRIT251 = 1	พันธุ์ RRIT251 = 7	พันธุ์ RRIT251 = 7
พันธุ์ RRIT226 = 2	พันธุ์ RRIT226 = 13	พันธุ์ RRIT226 = 1	พันธุ์ RRIT226 = 1	พันธุ์ RRIT226 = 13	พันธุ์ RRIT226 = 13
พันธุ์ BPM24 = 3	พันธุ์ BPM24 = 5	พันธุ์ BPM24 = 1	พันธุ์ BPM24 = 1	พันธุ์ BPM24 = 23	พันธุ์ BPM24 = 23
พันธุ์ RRM600 = 7	พันธุ์ RRM600 = 1	พันธุ์ RRM600 = 2	พันธุ์ RRM600 = 2	พันธุ์ RRM600 = 29	พันธุ์ RRM600 = 29
พันธุ์ PB235 = 9	พันธุ์ PB235 = 1	พันธุ์ PB235 = 1	พันธุ์ PB235 = 1	พันธุ์ PB235 = 8	พันธุ์ PB235 = 8
พันธุ์ RRIT251 = 4	พันธุ์ RRIT251 = 4	พันธุ์ RRIT251 = 24	พันธุ์ RRIT251 = 24	พันธุ์ RRIT251 = 1	พันธุ์ RRIT251 = 1
พันธุ์ RRIT226 = 1	พันธุ์ RRIT226 = 1	พันธุ์ RRIT226 = 28	พันธุ์ RRIT226 = 28	พันธุ์ RRIT226 = 1	พันธุ์ RRIT226 = 1
พันธุ์ BPM24 = 1	พันธุ์ BPM24 = 25	พันธุ์ BPM24 = 14	พันธุ์ BPM24 = 14	พันธุ์ BPM24 = 1	พันธุ์ BPM24 = 1
พันธุ์ RRM600 = 21	พันธุ์ RRM600 = 1	พันธุ์ RRM600 = 26	พันธุ์ RRM600 = 26	พันธุ์ RRM600 = 1	พันธุ์ RRM600 = 1
พันธุ์ PB235 = 6	พันธุ์ PB235 = 1	พันธุ์ PB235 = 8	พันธุ์ PB235 = 8	พันธุ์ PB235 = 1	พันธุ์ PB235 = 1
พันธุ์ RRIT251 = 1	พันธุ์ RRIT251 = 1	พันธุ์ RRIT251 = 1	พันธุ์ RRIT251 = 1	พันธุ์ RRIT251 = 1	พันธุ์ RRIT251 = 1
พันธุ์ RRIT226 = 4	พันธุ์ RRIT226 = 1	พันธุ์ RRIT226 = 1	พันธุ์ RRIT226 = 1	พันธุ์ RRIT226 = 1	พันธุ์ RRIT226 = 1
พันธุ์ BPM24 = 1	พันธุ์ BPM24 = 1	พันธุ์ BPM24 = 1	พันธุ์ BPM24 = 1	พันธุ์ BPM24 = 1	พันธุ์ BPM24 = 1
พันธุ์ RRM600 = 2	พันธุ์ RRM600 = 1	พันธุ์ RRM600 = 1	พันธุ์ RRM600 = 1	พันธุ์ RRM600 = 1	พันธุ์ RRM600 = 1
พันธุ์ PB235 = 15	พันธุ์ PB235 = 1	พันธุ์ PB235 = 1	พันธุ์ PB235 = 1	พันธุ์ PB235 = 1	พันธุ์ PB235 = 1

ภาพที่ 2: ภาพแสดงจำนวนของคุณลักษณะย่อยที่นับได้

หลังจากทำการนับจำนวนของคุณลักษณะย่อยแล้ว จะนำค่าที่นับได้ไปหารกับจำนวนของพันธุ์ยางพาราทั้งหมดที่มี ตามสูตรดังสมการที่ (2) ผลที่ได้จะเป็นค่าจำนวนที่นำไปใช้ทำตัวแบบการจำแนกดังภาพที่ 3

ลักษณะใบตัดตามขวาง	ลักษณะขอบใบ	ลักษณะรอยแผลก้านใบ	ลักษณะการท่ามพรหวงก้านใบ	ลักษณะฐานใบ	====พันธุ์====
พันธุ์ RRIT251 = 0.7419	พันธุ์ RRIT251 = 0.3226	พันธุ์ RRIT251 = 0.2581	พันธุ์ RRIT251 = 0.2581	พันธุ์ RRIT251 = 0.8065	พันธุ์ RRIT251 = 0.2000
พันธุ์ RRIT226 = 0.2258	พันธุ์ RRIT226 = 0.7742	พันธุ์ RRIT226 = 0.1290	พันธุ์ RRIT226 = 0.6452	พันธุ์ RRIT226 = 0.6129	พันธุ์ RRIT226 = 0.2000
พันธุ์ BPM24 = 0.0323	พันธุ์ BPM24 = 0.6129	พันธุ์ BPM24 = 0.1290	พันธุ์ BPM24 = 0.0000	พันธุ์ BPM24 = 0.6129	พันธุ์ BPM24 = 0.2000
พันธุ์ RRM600 = 0.8323	พันธุ์ RRM600 = 0.2258	พันธุ์ RRM600 = 0.8710	พันธุ์ RRM600 = 0.8710	พันธุ์ RRM600 = 0.2903	พันธุ์ RRM600 = 0.2000
พันธุ์ PB235 = 0.1935	พันธุ์ PB235 = 0.0968	พันธุ์ PB235 = 0.8065	พันธุ์ PB235 = 0.2903	พันธุ์ PB235 = 0.0968	พันธุ์ PB235 = 0.2000
พันธุ์ RRIT251 = 0.0645	พันธุ์ RRIT251 = 1.0000	พันธุ์ RRIT251 = 0.0323	พันธุ์ RRIT251 = 0.0323	พันธุ์ RRIT251 = 0.0968	พันธุ์ RRIT251 = 0.2000
พันธุ์ RRIT226 = 0.0968	พันธุ์ RRIT226 = 0.2258	พันธุ์ RRIT226 = 0.7742	พันธุ์ RRIT226 = 0.7742	พันธุ์ RRIT226 = 0.0968	พันธุ์ RRIT226 = 0.2000
พันธุ์ BPM24 = 0.0645	พันธุ์ BPM24 = 0.2258	พันธุ์ BPM24 = 0.0323	พันธุ์ BPM24 = 0.0323	พันธุ์ BPM24 = 0.7742	พันธุ์ BPM24 = 0.2000
พันธุ์ RRM600 = 0.0968	พันธุ์ RRM600 = 0.2258	พันธุ์ RRM600 = 0.0323	พันธุ์ RRM600 = 0.0323	พันธุ์ RRM600 = 0.7742	พันธุ์ RRM600 = 0.2000
พันธุ์ PB235 = 0.2903	พันธุ์ PB235 = 1.0000	พันธุ์ PB235 = 0.0323	พันธุ์ PB235 = 0.0323	พันธุ์ PB235 = 0.7742	พันธุ์ PB235 = 0.2000
พันธุ์ RRIT251 = 0.1290	พันธุ์ RRIT251 = 0.6129	พันธุ์ RRIT251 = 0.1613	พันธุ์ RRIT251 = 0.1613	พันธุ์ RRIT251 = 0.2258	พันธุ์ RRIT251 = 0.2000
พันธุ์ RRIT226 = 0.6774	พันธุ์ RRIT226 = 0.4194	พันธุ์ RRIT226 = 0.0323	พันธุ์ RRIT226 = 0.0323	พันธุ์ RRIT226 = 0.4194	พันธุ์ RRIT226 = 0.2000
พันธุ์ BPM24 = 0.0323	พันธุ์ BPM24 = 0.1935	พันธุ์ BPM24 = 0.1613	พันธุ์ BPM24 = 0.0323	พันธุ์ BPM24 = 0.7419	พันธุ์ BPM24 = 0.2000
พันธุ์ RRM600 = 0.7742	พันธุ์ RRM600 = 0.0323	พันธุ์ RRM600 = 0.9032	พันธุ์ RRM600 = 0.9032	พันธุ์ RRM600 = 0.9355	พันธุ์ RRM600 = 0.2000
พันธุ์ PB235 = 0.0968	พันธุ์ PB235 = 0.0323	พันธุ์ PB235 = 0.4516	พันธุ์ PB235 = 0.4516	พันธุ์ PB235 = 0.2581	พันธุ์ PB235 = 0.2000
พันธุ์ RRIT251 = 0.0323	พันธุ์ RRIT251 = 0.1290	พันธุ์ RRIT251 = 0.8387	พันธุ์ RRIT251 = 0.8387	พันธุ์ RRIT251 = 0.0323	พันธุ์ RRIT251 = 0.2000
พันธุ์ RRIT226 = 0.1290	พันธุ์ RRIT226 = 0.0323	พันธุ์ RRIT226 = 0.2581	พันธุ์ RRIT226 = 0.2581	พันธุ์ RRIT226 = 0.0323	พันธุ์ RRIT226 = 0.2000
พันธุ์ BPM24 = 0.0645	พันธุ์ BPM24 = 0.0323	พันธุ์ BPM24 = 0.0323	พันธุ์ BPM24 = 0.0323	พันธุ์ BPM24 = 0.0323	พันธุ์ BPM24 = 0.2000
พันธุ์ RRM600 = 0.4839	พันธุ์ RRM600 = 0.8065	พันธุ์ RRM600 = 0.0323	พันธุ์ RRM600 = 0.0323	พันธุ์ RRM600 = 0.0323	พันธุ์ RRM600 = 0.2000
พันธุ์ PB235 = 0.4839	พันธุ์ PB235 = 0.8065	พันธุ์ PB235 = 0.0323	พันธุ์ PB235 = 0.0323	พันธุ์ PB235 = 0.0323	พันธุ์ PB235 = 0.2000

ภาพที่ 3: แสดงค่าตัวแบบที่ได้จากการคำนวณเพื่อนำไปสร้างตัวแบบการจำแนกพันธุ์

4. ผลการดำเนินการวิจัย

การจำแนกพันธุ์ยางพาราจะใช้ตัวแบบการจำแนกพันธุ์ยางพาราโดยใช้ตัวจำแนกนออีพีเบย์ที่สร้างขึ้น ขั้นตอนการทำงานคือ ผู้ใช้งานต้องเลือกคุณลักษณะของพันธุ์ยางพาราให้ตรงกับลักษณะของพันธุ์ยางพาราที่ต้องการจำแนกพันธุ์ ดังภาพที่ 4

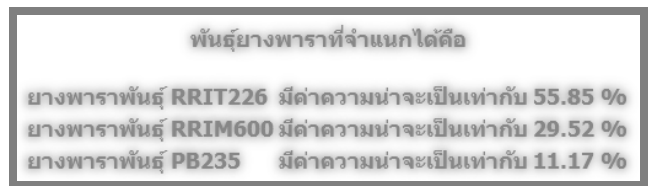
จำแนกพันธุ์ยางพารา

ลักษณะใบตัดตามขวาง :	▼ เว้าหรือรูปท้องเรือ ▼
ลักษณะขอบใบ :	▼ เรียบ ▼
ลักษณะรอยแผลก้านใบ :	▼ รูปหัวใจ ▼
ลักษณะความกว้างของฉัตร :	▼ แคบ ▼
ลักษณะความเป็นมันของใบ :	▼ เป็นมัน ▼
ลักษณะผิวใบ/แผ่นใบ :	▼ เรียบ ▼
ลักษณะความยาวของก้านใบ :	▼ ยาว ▼
ลักษณะรูปร่างของฐานก้านใบ :	▼ ฐานมีร่อง ▼
ลักษณะความยาวก้านใบย่อย :	▼ สั้น ▼
ลักษณะเปลือกส่วนสีเขียว :	▼ รุหายใจไม่ชัด ▼

คำนวณ

ภาพที่ 4: การเลือกลักษณะของพันธุ์ยางพาราเพื่อจำแนกพันธุ์

จากนั้นระบบจะทำการนำค่าที่ผู้ใช้งานได้เลือก ไปเปรียบเทียบกับตัวแบบการจำแนกพันธุ์ที่ได้สร้างไว้ โดยจะคำนวณตามสูตรดังสมการที่ (2) ผลลัพธ์ที่ได้จะอยู่ในรูปเปอร์เซ็นต์ ค่าเปอร์เซ็นต์ที่แสดงจะบอกถึงความน่าจะเป็นของต้นยางพาราที่นำมาจำแนกพันธุ์ว่าเป็นพันธุ์อะไร และความเป็นไปได้ที่ต้นยางพาราจะเป็นพันธุ์อื่นที่มีเปอร์เซ็นต์ ดังภาพที่ 5



ภาพที่ 5: ผลลัพธ์ที่ได้จากการจำแนกพันธุ์ยางพารา

จากการทดลองจำแนกพันธุ์ยางพารา ตามข้อมูลของ dataset ที่มี โดยได้ทำการเลือกคุณลักษณะของพันธุ์ยางพาราตามภาพที่ 4 ประกอบไปด้วย ใบตัดตามขวางเลือก “เว้าหรือรูปท้องเรือ”, ขอบใบเลือก “เรียบ”, รอยแผลก้านใบเลือก “รูปหัวใจ”, ความกว้างของฉัตรเลือก “แคบ”, ความเป็นมันเลือก “เป็นมัน”, ผิวใบเลือก “เรียบ”, ความยาวของก้านใบเลือก “ยาว”, รูปร่างของฐานก้านใบเลือก “ฐานมีร่อง”, ความยาวก้านใบย่อยเลือก “สั้น” และเปลือกส่วนสีเขียวเลือก “รุหายใจไม่ชัด” ผลลัพธ์ที่ได้คือ เป็นยางพาราพันธุ์ RRIT 226 มีค่าความน่าจะเป็นเท่ากับ 55.85%, ยางพาราพันธุ์ RRM 600 มีค่าความน่าจะเป็นเท่ากับ 29.52% และยางพาราพันธุ์ PB 235 มีค่าความน่าจะเป็นเท่ากับ 11.17%

5. บทสรุปและข้อเสนอแนะ

การวิจัยในครั้งนี้ใช้เทคนิคนาอิวเบย์ (Naive Bayes) ในการสร้างตัวแบบการจำแนกพันธุ์ยางพารา ที่นำไปใช้ในการจำแนกพันธุ์ยางพาราที่ยังไม่ทราบพันธุ์ แต่ทราบรายละเอียดต่างๆ เกี่ยวกับลักษณะเฉพาะประจำพันธุ์ เช่น จักร ใบ ก้านใบ ก้านใบย่อย เปลือก และตาของต้นยางพารา และมีการใช้เทคนิค Spearman correlation coefficient และ Chi-square ในการเลือกคุณลักษณะเด่นจากคุณลักษณะย่อยทั้งหมด 34 คุณลักษณะ ทำให้ได้คุณลักษณะเด่นทั้งหมด 12 คุณลักษณะ จากนั้นหาคุณลักษณะเด่นที่มีค่าความถูกต้องสูงสุด มาใช้ในกระบวนการสร้างตัวแบบการจำแนกพันธุ์ ทำให้ได้คุณลักษณะเด่นที่ดีที่สุด 10 คุณลักษณะ ผลที่ได้จากการจำแนกพันธุ์ยางพาราด้วยตัวแบบการจำแนกนาอิวเบย์มีความค่าความถูกต้องเท่ากับ 98.02% และได้ทำการเปรียบเทียบความถูกต้องในการจำแนกพันธุ์ยางพาราโดยผู้เชี่ยวชาญ มีค่าความถูกต้อง 83.33 %

จากผลการเปรียบเทียบที่ได้ สรุปได้ว่าตัวแบบการจำแนกพันธุ์ยางพาราโดยใช้ตัวจำแนกนาอิวเบย์ สามารถจำแนกพันธุ์ยางพาราได้ถูกต้อง 98.02% ซึ่งมีความแม่นยำอยู่ในระดับสูง และผลการจำแนกของผู้เชี่ยวชาญมีความถูกต้อง 83.33%

การจำแนกพันธุ์ยางพาราที่ถูกต้องจะช่วยให้เกษตรกรทราบพันธุ์ยางพาราที่ถูกต้องจากการนำเอาคุณลักษณะของต้นยางมาจำแนก หลักการนี้สามารถนำไปพัฒนาต่อยอดให้สามารถจำแนกพันธุ์ยางพาราได้มากกว่า 5 พันธุ์ และสามารถนำไปพัฒนาเป็นโปรแกรมคอมพิวเตอร์ ใช้ในหน่วยงานสถาบันวิจัยยางพาราของภาครัฐและเอกชนได้

เอกสารอ้างอิง

- [1] สถาบันวิจัยยาง กรมวิชาการเกษตร. [Online]. แหล่งที่มา : <http://www.live-rubber.com/index.php/para-rubber-articles/42-para-rubber-clones/201-rubber-clone-identification>
- [2] วรรัตน์ รุ่งรวุฒิ. [ม.ป.ป.]. “การทำเหมืองข้อมูล Data Mining (การจำแนกประเภทโดยใช้กฎของเบย์ Bayesian Learning).” ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น.
- [3] นิเวศ จิระวิชิตชัย และนรินทร์ พนาวาส. “การจำแนกความคิดเห็นโดยใช้เทคนิคการเรียนรู้ของเครื่อง.” สาขาวิชาคอมพิวเตอร์ธุรกิจ คณะเทคโนโลยีสารสนเทศมหาวิทยาลัยศรีปทุม วิทยาเขตชลบุรี. 2554.
- [4] จูติมา เกษมศรีธนาวัฒน์ และธนัสณี เพียรตระกูล. “การจำแนกความคิดเห็นโดยใช้ตัวจำแนกแบบเบย์ร่วมกับการเลือกคุณลักษณะด้วยอัลกอริทึมวิธีลิฟ.” สาขาวิชาเทคโนโลยีการจัดการระบบสารสนเทศ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยมหิดล. 2554.
- [5] บุญเสริม กิจศิริกุล. “ปัญญาประดิษฐ์.” ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย. 2548.
- [6] หทัยชนก กรสี. “การคำนวณน้ำหนักของคุณลักษณะข้อมูลสำหรับตัวจำแนกในอิวเบย์โดยวิธีการจัดกลุ่มข้อมูลและต้นไม้ตัดสินใจ.” วิทยาการคอมพิวเตอร์ บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์. 2551.
- [7] Choochart Haruechaiyasak. “A Tutorial on Naive Bayes Classification.” 2008.
- [8] สถิติและการวิเคราะห์ข้อมูลทางสุขภาพ. “บทที่ 7 สหสัมพันธ์ (Correlation)” [Online]. แหล่งที่มา : http://intraserver.nurse.cmu.ac.th/mis/download/course/lec_567730_lesson_07.pdf
- [9] จันทิมา หัสสินี และนุริสสา อูมาหมั่น. “นอนพาราเมตริก” [Online]. แหล่งที่มา : www.saruthipong.com/port/document/299-705/299-705-13.ppt
- [10] สถาบันวิจัยยาง กรมวิชาการเกษตร. *ฐานข้อมูลเชื้อพันธุ์พืช : พันธุ์ยางพารา*. กรุงเทพฯ: ชุมชุมสหกรณ์การเกษตรแห่งประเทศไทย จำกัด. 2554.
- [11] สถาบันวิจัยยาง กรมวิชาการเกษตร. *คำแนะนำการปลูกยางพาราในพื้นที่ปลูกยางใหม่*. พิมพ์ครั้งที่ 7. กรุงเทพฯ: ชุมชุมสหกรณ์การเกษตรแห่งประเทศไทย. 2553.
- [12] ปาริชาติ เจริญสินธุ์ และศุภเศรษฐ์ วิริยะพิทักษ์. “คู่มือการใช้ WEKA : Classification ใช้ตัวอย่าง weather.” ภาควิชาระบบสารสนเทศทางคอมพิวเตอร์ คณะบริหารธุรกิจ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี. 2553.

โมเดลการวิเคราะห์โรคในสุกร โดยใช้โครงข่ายประสาทเทียม

A Model for Swine Disease Analysis using Neural Network

จตุรรัตน์ ตั้งกิตติวัฒน์ (Jutarat Tangkittiwat)¹ และ นลินภัทร์ พรวัฒน์ปรียกอร์ (Nalinpat Porrawatpreyakorn)²

^{1,2}ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
palm.tangkittiwat@gmail.com, nalinpatp@kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาโมเดลการวิเคราะห์โรคในสุกร โดยใช้โครงข่ายประสาทเทียม โปรแกรม WEKA ถูกนำมาใช้เป็นเครื่องมือในการทดสอบการวิเคราะห์และจำแนกข้อมูล โดยตัวแปรในการจำแนกข้อมูลนั้น เป็นตัวแปรที่สำคัญและเกี่ยวข้องกับสุกร เช่น เพศ อายุ น้ำหนักและอาการของโรค โดยเก็บข้อมูลจากสุกรที่เป็นโรค ระหว่างปี 2553 - 2556

ผลจากการทดสอบโมเดลพบว่ามีค่าความถูกต้องสูงถึง 99.5% ซึ่งอยู่ในระดับดีมาก ดังนั้นผลที่ได้จากโมเดลนี้จะถูกนำไปพัฒนาต่อให้อยู่ในรูปแบบของระบบสารสนเทศเพื่อให้ง่ายต่อการใช้งาน ทั้งนี้ระบบดังกล่าวจะนำเสนอวิธีการรักษาและคำแนะนำเพื่อลดการสูญเสียในอาชีพเกษตรกร อีกทั้งสนับสนุนการพัฒนาองค์ความรู้เกี่ยวกับโรคในสุกรที่ติดต่อมาสู่คนอีกด้วย

คำสำคัญ: โครงข่ายประสาทเทียม การวิเคราะห์โรค

การจำแนกข้อมูล

Abstract

This research aims to develop a model for swine disease analysis using neural network. WEKA was used as a tool for data analysis and classification. Related variables used for the classification include gender, age, weight, and symptoms by collecting data from diseased pigs during the years 2554-2556. The model test result reveals 99.5% accuracy. In the near future, this results will be used for developing an information system which aims at offering guidelines in and recommendations on treatment of swine diseases in order to reduce any loss,

as well as providing the support to build knowledge about pig-to-human-contact diseases.

Keyword: Artificial Neural Network

Analysis of Disease, Data classification

1. บทนำ

ในปัจจุบันอุตสาหกรรมการเลี้ยงสุกรในประเทศไทยได้พัฒนาเจริญก้าวหน้าขึ้นเป็นอย่างมาก เนื่องจากราคาจำหน่ายเนื้อสุกรในท้องตลาดมีราคาสูงขึ้น จึงเป็นสิ่งจูงใจให้ประชากรหันมาเลี้ยงสุกรกันมากขึ้น ซึ่งการเลี้ยงสุกรจำนวนมากก็ย่อมประสบปัญหาเกี่ยวกับโรคของสุกรที่ผู้เลี้ยงไม่สามารถหลีกเลี่ยงได้ โดยโรคที่เกิดขึ้นกับสุกรนั้นมีทั้งโรคระบาดและโรคไม่ระบาด อีกทั้งในปัจจุบันโรคได้มีวิวัฒนาการมากขึ้นโดยสามารถแพร่กระจายไปสู่คนทำให้เป็นอันตรายต่อชีวิตได้ [1], [2] ในปัจจุบันเทคโนโลยีสารสนเทศได้มีความเจริญก้าวหน้ามากขึ้นและได้เข้ามามีบทบาทในชีวิตประจำวันเพิ่มขึ้น เพียงแค่มี Smartphone หรืออุปกรณ์ด้านอิเล็กทรอนิกส์ต่างๆ ก็สามารถรับข้อมูลข่าวสารสารสนเทศได้อย่างรวดเร็ว ซึ่งธุรกิจด้านเกษตรกรรมก็มีความจำเป็นต้องอาศัยเทคโนโลยีสารสนเทศเช่นกัน โดยนำมาช่วยสนับสนุนการตัดสินใจให้เครื่องคอมพิวเตอร์ไปแก้ปัญหา ช่วยในการจัดระบบองค์ความรู้ให้กับเกษตรกรผู้เลี้ยงสุกรหรือบุคคลอื่น ๆ ที่สนใจ ให้ตระหนักถึงความสำคัญในการจัดการระบบการเลี้ยงที่เป็นมาตรฐานและควรปฏิบัติต่อสุกรอย่างถูกต้อง [3] เพราะถ้าสุกรมีสุขภาพดีก็จะทำให้ผู้เลี้ยงสามารถลดต้นทุนค่ายาและค่ารักษา ลงได้ ทำให้สุกรสามารถให้ผลผลิตได้เต็มความสามารถและผลตอบแทนสูง [1]

การพัฒนาเครื่องมือช่วยในการตรวจวิเคราะห์โรค จำเป็นต้องอาศัยความรู้ ความสามารถในการตัดสินใจที่มีความยุ่งยากและซับซ้อนซึ่งเป็นการยากที่จะจำแนกค่าต่าง ๆ ได้อย่างถูกต้อง หากไม่ได้มีความชำนาญเฉพาะด้านอย่างแท้จริงก็อาจจะเกิดการผิดพลาดได้ง่าย ดังนั้นผู้วิจัยจึงได้ศึกษาเทคนิคโครงข่ายประสาทเทียม (Neural Network) ซึ่งเป็นวิธีการทางด้านปัญญาประดิษฐ์เพื่อนำความสามารถของโครงข่ายประสาทเทียมเข้ามาช่วยในการวิเคราะห์โรค โดยวิเคราะห์ปัจจัยจาก เพศ อายุ น้ำหนักและอาการที่สังเกตได้ ซึ่งจะทำให้ผู้ใช้สามารถวิเคราะห์ความเสี่ยงได้ด้วยตนเองเบื้องต้นว่าแต่ละอาการที่เกิดในสุกรมี่ความน่าจะเป็นโรคอะไร และมีโรคใดบ้างในสุกรที่สามารถติดต่อมาสู่คนได้ พร้อมทั้งให้คำแนะนำอธิบายวิธีการรักษา [4] เพื่อเป็นประโยชน์กับผู้ต่อไป

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 โรคสุกร การเปลี่ยนแปลงใด ๆ ของสุกรที่ผิดปกติโดยอาการอาจเกิดขึ้นอย่างต่อเนื่อง โดยมีเชื้อโรคหรือปัจจัยอื่น ๆ เป็นสาเหตุอันเป็นผลทำให้สุกรเสียชีวิตหรือเจริญเติบโตได้ช้าลง โดยมีทั้งทั้งโรคระบาด คือโรคที่สามารถแพร่จากสุกรป่วยตัวหนึ่งไปยังสุกรตัวอื่นได้อย่างรวดเร็วและโรคไม่ระบาด คือโรคที่ไม่สามารถแพร่จากสุกรป่วยตัวหนึ่งไปยังสุกรตัวอื่น ๆ หรือเป็นโรคเฉพาะตัว [2]

2.2 การวิเคราะห์โรค เป็นการระบุว่าโรคอะไรโดยอาศัยข้อมูลบางส่วนหรือทั้งหมดของลักษณะอาการผิดปกติ ซึ่งจะสามารถระบุได้ว่าโรคหรือใกล้เคียงกับโรคอะไร และเกิดจากสาเหตุอะไร อีกทั้งควรรักษาอย่างไรเพื่อหาแนวทางป้องกันที่ดีที่สุด [5]

2.3 กลไกการวินิจฉัย หมายถึง ส่วนจำลองกระบวนการความคิดของมนุษย์ โดยใช้ข้อมูลที่ป้อนเข้ามาประมวลกับกฎข้อเท็จจริงที่มีอยู่ในระบบ แล้วสรุปผลเป็นข้อเท็จจริงใหม่ซึ่งจะถูกนำเสนอสู่ผู้ใช้งานทางส่วนติดต่อกับผู้ใช้งาน [3]

2.4 โครงข่ายประสาทเทียม (Artificial Neural Network: ANN) เป็นแบบจำลองทางคณิตศาสตร์ที่ถูกสร้างขึ้นมาเพื่อจำลองการทำงานให้คล้ายคลึงกับการทำงานของระบบประสาทในสมองของมนุษย์ โดยมีวัตถุประสงค์เพื่อให้

ระบบสามารถรู้จำรูปแบบ (Pattern Recognition) และนำมาสร้างองค์ความรู้ใหม่ได้ (Knowledge Extraction) เป็นต้น [6]

โครงข่ายประสาทเทียมแบบ Multi-layer Perceptron (MLP) เป็นรูปแบบหนึ่งของโครงข่ายประสาทเทียมที่มีโครงสร้างเป็นแบบชั้นใช้สำหรับงานที่มีความซับซ้อนได้ผลเป็นอย่างดี โดยมีกระบวนการฝึกฝนเป็นแบบการเรียนรู้แบบมีการควบคุม (Supervised Learning) และใช้ขั้นตอนการส่งค่าย้อนกลับ (Backpropagation) สำหรับการฝึกฝนกระบวนการเรียนรู้ [4]

การทำงานของโครงข่ายประสาทเทียมเมื่อมีข้อมูลนำเข้า (Input) เข้ามายังโครงข่ายประสาท (Network) ก็นำ Input มาคูณกับค่าน้ำหนัก (Weight) ของแต่ละขา ผลที่ได้จาก Input ทุก ๆ ขาของ Neuron จะนำมาบวกกันแล้วก็นำมาเทียบกับระดับการกระตุ้นต่ำสุดที่ทำให้เกิดการตอบสนอง (Threshold) ที่กำหนดไว้ ถ้าผลรวมมีค่ามากกว่า Threshold แล้ว Neuron จะส่งผลลัพธ์ (Output) ออกไป และ Output ก็จะถูกส่งไปยัง Input ของ Neuron อื่น ๆ ที่เชื่อมกันใน Network จนได้เป็นผลลัพธ์ที่ต้องการ

หลังจากนั้นจะต้องมีการตรวจสอบผลลัพธ์ที่ได้จากระบบว่ามีความคลาดเคลื่อนเพียงใด หากความคลาดเคลื่อนจากเป้าหมายมากเกินไปจะต้องมีการนำค่าความคลาดเคลื่อนนี้ไปปรับน้ำหนักการเรียนรู้ใหม่ (Weight) แบบแพร่กระจายย้อนกลับ (Backpropagation) ซึ่งจะเป็นการปรับน้ำหนักความคลาดเคลื่อนจากชั้นผลลัพธ์ไปยังชั้นก่อนหน้า และทำการปรับน้ำหนักไปเรื่อย ๆ จนกระทั่งถึงชั้นรับข้อมูล ซึ่งกระบวนการเรียนรู้แบบนี้จะต้องอาศัยการทำซ้ำ หลายรอบจนกว่าจะได้ผลลัพธ์ตามที่กำหนดหรือได้ค่าความคลาดเคลื่อนที่น้อยจนพอยอมรับได้ ซึ่งจำนวนรอบนี้จะอยู่กับความยากง่ายของปัญหา ขนาดข้อมูล รวมไปถึงจำนวนชั้นของโครงสร้างของโครงข่ายที่เราสร้างไว้ด้วย [8] ในการหาค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Squared Error: MSE) เราสามารถวิเคราะห์ค่าความผิดพลาดของการจำแนกประเภท (Classify) ในรายละเอียดของข้อมูลได้มากขึ้นเพราะเป็นค่าความผิดพลาดที่เกิดขึ้นจริงเฉลี่ยจากข้อมูลทั้งหมด มีประโยชน์สำหรับนำไปใช้งานจริง ซึ่งค่า MSE ยิ่งน้อย หมายถึง การพยากรณ์ยิ่งแม่นยำ

2.5 งานวิจัยที่เกี่ยวข้อง

งานวิจัยหรือเอกสารอื่น ๆ ที่เกี่ยวกับข้องกับระบบการวิเคราะห์โรคหรือการจำแนกโรค พบว่าอัลกอริทึมที่นำมาช่วยในการวิจัยนั้นมีหลายด้าน โดยการเลือกใช้อัลกอริทึมใดนั้นขึ้นอยู่กับผู้วิจัยแต่ละท่านว่าข้อมูลที่ใช้ในการทดลองเป็นข้อมูลประเภทใด ขนาดเท่าใด ใช้ในการแก้ปัญหาของงานแบบใด ซึ่งอัลกอริทึมโครงข่ายประสาทเทียมก็เป็นอัลกอริทึมหนึ่งที่มีผู้วิจัยเลือกใช้ในการวิเคราะห์โรค เพราะว่าเป็นอัลกอริทึมที่มีความสามารถหลากหลายมาก ทั้งการพยากรณ์ การจัดกลุ่ม การแบ่งกลุ่ม ฯลฯ ตัวอย่าง เช่น การใช้โครงข่ายประสาทเทียมเพื่อเช้ปตรอนหลายชั้นจำแนกประเภทของมะเร็งเม็ดเลือดขาว แล้วนำมาเปรียบเทียบกับอัลกอริทึมต้นไม้ตัดสินใจและอัลกอริทึมเนอโฟเบย์ ผลที่ได้คือ การจำแนกประเภทของมะเร็งเม็ดเลือดขาวโดยใช้โครงข่ายประสาทเทียมเพื่อเช้ปตรอนหลายชั้นให้ความถูกต้องมากที่สุด รองลงมาคือเนอโฟเบย์และอัลกอริทึมต้นไม้ตัดสินใจ [9] หรือจะเป็นการนำเสนอระบบการพยากรณ์จำนวนผู้ป่วยโรคหัวใจ โดยใช้อัลกอริทึมโครงข่ายประสาทเทียม ผลที่ได้พบว่าสามารถพยากรณ์ค่าล่วงหน้าได้ 1 ปี อีกทั้งสามารถนำไปประยุกต์ใช้งานได้จริงโดยวัดจากคุณภาพของความผิดพลาดของการพยากรณ์ [7]

3. วิธีการดำเนินการ

3.1 การศึกษาข้อมูล

ทำการศึกษาเอกสาร ข้อมูลต่าง ๆ ที่เกี่ยวกับโรคที่เกิดในสุกรและที่มีโอกาสติดต่อมาสู่คน ศึกษาถึงตัวแปรที่มีผลทำให้เกิดโรคในสุกรจากหนังสืองานวิจัย เครือข่ายอินเทอร์เน็ต รวมถึงบุคคลที่มีความเชี่ยวชาญด้านโรคในสุกร เช่น สัตว์แพทย์ สัตว์บาล เพื่อนำมาจำแนกอาการนั้นว่าเป็นโรคอะไรหรือมีอาการกลุ่มใดเป็นกลุ่มเสี่ยงที่จะทำให้อาจเกิดโรคนี้ได้ โดยผู้วิจัยได้นำข้อมูลตัวอย่างสุกรที่เกิดโรคทั้ง 22 ชนิด มาจากบริษัทแห่งหนึ่งที่ได้ดูแลเรื่องสารเสริมและผลิตภัณฑ์รักษาโรคสุกร ซึ่งเน้นเรื่องของการดูแลสุขภาพสัตว์ให้แข็งแรงเพื่อให้ผลผลิตที่เต็มกับผู้เลี้ยง โดยเป็นข้อมูลระหว่างช่วงปี พ.ศ. 2553 – 2556 ซึ่งข้อมูลที่ได้จะนำมาทำการวิเคราะห์โดยใช้อัลกอริทึมของโครงข่ายประสาทเทียมแบบ Multi-layer

Perceptron (MLP) เพื่อทำการหาประสิทธิภาพเพิ่มความแม่นยำในการทำงาน

3.2 การเตรียมข้อมูล

ขั้นตอนการเตรียมข้อมูลก่อนที่จะส่งไปเข้าสู่กระบวนการหาอัลกอริทึมและวิเคราะห์ ผู้วิจัยได้นำข้อมูลแอททริบิวต์ เพศ แบ่งเป็น ตัวผู้ และตัวเมีย แอททริบิวต์อายุ แบ่งเป็นช่วงอายุ 3 ช่วง คือ ช่วงต่ำกว่าหรือเท่ากับ 3 เดือน (< 3 เดือน) ช่วง 4 – 12 เดือน และ ช่วง 1 ปีขึ้นไป แอททริบิวต์น้ำหนักแบ่งช่วงน้ำหนัก 3 ช่วง คือ ช่วงต่ำกว่าหรือเท่ากับ 60 กิโลกรัม (< 60 กิโลกรัม) ช่วง 61 – 150 กิโลกรัม และ ช่วง 151 กิโลกรัมขึ้นไป และ แอททริบิวต์อาการได้แบ่งเป็น 34 อาการ จากข้อมูลโรค 22 ชนิดโรค และแอททริบิวต์เป้าหมายก็จะเป็นตัวระบุว่าสุกรเป็นโรคใด แสดงรายละเอียดดังตารางที่ 1

ตารางที่ 1: รายละเอียดของแอททริบิวต์

แอททริบิวต์ที่	ชื่อ	รายละเอียด	ชนิด
1	SEX	0= ตัวเมีย 1= ตัวผู้	ตัวเลข
2	AGE	0= < 3 เดือน 1= 4-12 เดือน 2 = 1 ปีขึ้นไป	ตัวเลข
3	Weight	0 = <60 กิโลกรัม 1 = 61-150 กิโล 2 = 151 กิโลขึ้นไป	ตัวเลข
4 – 37	Symptom 1- 34	0 = ไม่มีอาการ 1= มีอาการ	ตัวเลข
38	Name_ Diseases	ชื่อของโรคที่เกิดตั้งแต่ A1 – A22	ตัวเลข

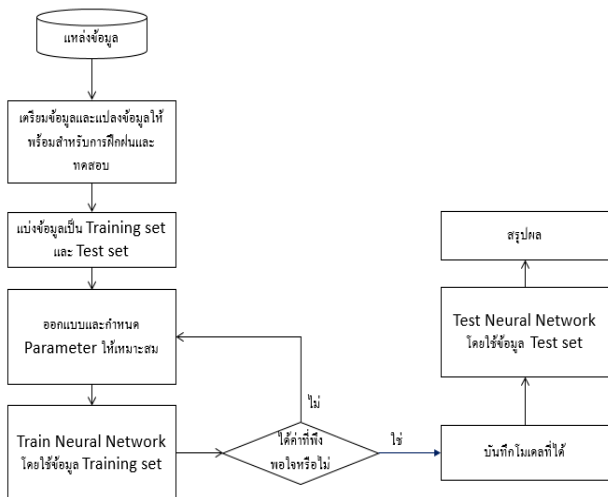
เพศ บางโรคมักเลือกเพศ เช่น โรคแท้งติดต่อหรือบรูเซลโลซิส จะเป็นเฉพาะในตัวเมีย หรือโรคที่อกโซพลาสโมซิสโรคจะมีโอกาสเกิดในเพศเมียมากกว่าเพศผู้ เราไม่สามารถแก้ไขเพศได้ จึงต้องระวังให้มากขึ้น [2]

อายุบางช่วงอายุมีความเสี่ยงต่อการเกิดโรคมกกว่าวัยอื่น ๆ เช่น ลูกสุกรมักจะยังไม่มิกุมิคุ้มกันที่ดี ทำให้เสี่ยงต่อการเป็นโรคต่าง ๆ มากกว่าสุกรขุน เราแก้ไขอายุไม่ได้แต่ควรเพิ่มวัคซีนป้องกันโรคหรือมีการจัดการสุขภาพที่ดี ก็จะลดความ

เสี่ยงได้ [2] และอาการ เป็นตัวบ่งชี้ว่า อาการแบบใดมีความเสี่ยงจะเป็นโรคใดบ้าง

3.3 การพัฒนาโมเดลและการสอนข้อมูล

การสร้างแบบจำลองและการสอนข้อมูล ผู้วิจัยได้พัฒนาแบบจำลองด้วยโปรแกรมการทำเหมืองข้อมูลเวกา (WEKA) รุ่น 3.6.9 ซึ่งเป็นโปรแกรมสำเร็จรูปถูกพัฒนาขึ้นด้วยภาษาจาวา โดยเลือกใช้งานแบบ Classify จากนั้นเลือกใช้ Function “Multi-layer Perceptron (Backpropagation)” เพื่อหาค่าน้ำหนัก ค่าไบแอสที่เหมาะสม และความคลาดเคลื่อนที่ต่ำที่สุด โดยข้อมูลที่น่ามาวิจัยนี้คือข้อมูลของสุกรที่เกิดโรค ซึ่งข้อมูลชุดนี้ได้แยกโรคในสุกรออกเป็น 22 คลาส (22 โรค) ข้อมูลแต่ละชุดประกอบด้วยแอททริบิวต์ที่ระบุถึงปัจจัยที่เกี่ยวข้องกับการจำแนกข้อมูล (ข้อมูลอินพุต) 37 แอททริบิวต์ โดยแอททริบิวต์ที่ 38 จะแสดงคลาสของโรคในสุกร ชุดข้อมูลที่ใช้ทั้งหมดมี 2,000 ชุด ได้ถูกแบ่งออกเป็นข้อมูลสำหรับปรับสอน (Training Set) จำนวน 80 % คือ 1,600 ชุด และข้อมูลสำหรับการทดสอบ (Testing Set) จำนวน 20 % คือ 400 ชุด โดยการวัดประสิทธิภาพของงานวิจัยนี้ วัดค่าความถูกต้องของการจัดกลุ่มข้อมูล และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย MSE ที่มีค่าน้อยที่สุด จากการปรับเปลี่ยนค่า Parameter ให้เหมาะสม โดยมีกระบวนการทำงาน แสดงดังภาพที่ 1



ภาพที่ 1: กระบวนการทำงานภาพรวมระบบ

4. ผลการดำเนินงาน

การจำแนกโดยใช้เทคนิคคาด้าไมน์นิงด้วยอัลกอริทึม MLP ในส่วนนี้จะเป็นการนำโครงสร้างของ Multi-layer

Perceptron เข้ามาทดสอบเพื่อสร้างโมเดลวิเคราะห์โรคในสุกร โดยใช้โครงข่ายประสาทเทียม เพื่อนำผลที่ได้ไปใช้ในการพัฒนาระบบวิเคราะห์โรคในสุกร โดยใช้โครงข่ายประสาทเทียมที่ให้ค่าความถูกต้องมากที่สุด และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย MSE น้อยที่สุด

ผู้วิจัยได้ทำการปรับเปลี่ยน Node ใน Hidden Layer โดยเริ่มตั้งแต่ 1 ไปจนถึง 9 โดยใช้อัตราการเรียนรู้ (Learning Rate) 0.3 ค่าโมเมนตัม (Momentum) 0.2 และรอบการเรียนรู้ (Training Time) 500 รอบ เพื่อหาจำนวน Hidden Layer ที่เหมาะสมมากที่สุด ซึ่งผลลัพธ์ที่ได้ แสดงดังตารางที่ 2

ตารางที่ 2: การปรับค่า Hidden Layer

Hidden Layer	Accuracy	MSE
1	57.375%	0.01086336
2	69.25%	0.01747684
3	87.125%	0.00956484
4	97.0625%	0.00289444
5	99%	0.00099856
6	99.8125%	0.00026244
7	99.8125%	0.00020164
8	99.875%	0.00013924
9	99.875%	0.00012544

จากตารางที่ 2 ได้ทำการปรับค่า Hidden Layer เพื่อให้ได้ค่าที่ถูกต้องมากที่สุด และให้มีค่าความผิดพลาดที่น้อยที่สุด โดยจะเห็นว่าผลการปรับค่า Hidden Layer ที่ดีที่สุดจะมีทั้ง Hidden Layer ที่ 8 และ 9 ที่มีค่าความถูกต้องเท่ากัน แต่ถ้าเปรียบเทียบถึงค่าความคลาดเคลื่อนกำลังสองเฉลี่ย MSE จะเห็นได้ว่าการปรับ Hidden Layer ที่ 9 จะได้ผลดีที่สุด

หลังจากได้ค่า Hidden Layer ที่ให้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยที่น้อยที่สุดแล้ว ทางผู้วิจัยจึงได้เลือกทำการปรับเปลี่ยน Node ใน Hidden Layer เท่ากับ 9 และทำการปรับอัตราการเรียนรู้ (Learning Rate) 0.3, 0.5, 0.7 และ 0.9 ตามลำดับ ค่าโมเมนตัม (Momentum) 0.2, 0.4, 0.6 และ 0.8 ตามลำดับ และรอบการเรียนรู้ (Training Time) 500 รอบ เพื่อหาโมเดลโครงข่ายประสาทเทียมที่ดีที่สุด ซึ่งผลลัพธ์ที่ได้ แสดงดังตารางที่ 3

ตารางที่ 3: เปรียบเทียบประสิทธิภาพของ MLP

Learning Rate	Momentum	Accuracy	MSE
0.3	0.2	99.875%	0.00012544
0.3	0.4	99.875%	0.00011025
0.3	0.6	99.875%	0.00009801
0.3	0.8	99.875%	0.00009604
0.5	0.2	99.875%	0.00010404
0.5	0.4	99.875%	0.00009604
0.5	0.6	99.875%	0.00009801
0.5	0.8	97%	0.00194481
0.7	0.2	99.875%	0.00009604
0.7	0.4	99.875%	0.00009025
0.7	0.6	99.875%	0.00008649
0.7	0.8	70.5625%	0.014884
0.9	0.2	99.875%	0.00009216
0.9	0.4	99.875%	0.00008836
0.9	0.6	99.75%	0.00020736
0.9	0.8	47%	0.02499561

จากตารางที่ 3 ได้ทำการปรับค่าพารามิเตอร์ เพื่อให้ได้ค่าความถูกต้องมากที่สุด และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย น้อยที่สุด โดยผลการปรับค่าพารามิเตอร์ที่ดีที่สุด คือ ค่าอัตราการเรียนรู้ (Learning Rate) จะอยู่ที่ระดับ 0.7 และค่าโมเมนตัม (Momentum) อยู่ที่ 0.6 โดยผลลัพธ์ที่ได้ให้ค่าความถูกต้องถึง 99.875% และมีค่าความคลาดเคลื่อนกำลังสองเฉลี่ย MSE เท่ากับ 0.00008649 ซึ่งถือว่าอยู่ในระดับดีมาก

หลังจากทำการสอนข้อมูลจนได้โมเดลโครงข่ายประสาทเทียมที่ดีที่สุดแล้ว จึงนำข้อมูลที่ได้แบ่งไว้สำหรับทดสอบจำนวน 400 ชุด มาทดสอบโมเดลโครงข่ายประสาทเทียมที่ผ่านการเรียนรู้มาแล้ว เพื่อการวัดประสิทธิภาพ (Evaluation) โมเดลโครงข่ายประสาทเทียม โดยผลลัพธ์ที่ได้ให้ค่าความถูกต้องถึง 99.5% และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย MSE เท่ากับ 0.0000256 ซึ่งถือว่าอยู่ในระดับดีมาก สามารถนำไปใช้พัฒนาต่อให้อยู่ในรูปแบบของระบบสารสนเทศเพื่อวิเคราะห์โรคในสุกรได้จริงและช่วยเพิ่มประสิทธิภาพความน่าเชื่อถือของข้อมูล

5. สรุปและข้อเสนอแนะ

5.1 สรุปผลการวิจัย

ผู้วิจัยได้นำข้อมูลที่เตรียมไว้สำหรับการเรียนรู้ไปทำการทดสอบกับอัลกอริทึม J.48 Decision Tree, RBFNetwork และ Multi-layer Perceptron (MLP) โดยใช้ (Classifier) เป็นตัวคัดแยก เพื่อเปรียบเทียบว่าอัลกอริทึมใดเหมาะที่จะนำมาใช้งานกับข้อมูลชุดนี้มากที่สุด ซึ่งผลลัพธ์ที่ได้จากการทดสอบแสดงดังตารางที่ 4

ตารางที่ 4: การเปรียบเทียบค่า MSE และค่าความถูกต้องของ Classified Instance

สิ่งที่เปรียบเทียบ	J48 Decision Tree	RBF Network	Multi-layer Perceptron (MLP)
Correctly Classified Instances	99.5%	98.6875%	99.8125%
Incorrectly Classified Instances	0.5	0.3125%	0.125%
MSE	0.0004040	0.0001612	0.0000864

จากตารางที่ 4 แสดงให้เห็นว่าผลลัพธ์ของทุกอัลกอริทึมมีค่าความถูกต้องและค่าความผิดพลาด MSE ที่ใกล้เคียงกัน แต่อัลกอริทึมที่ให้ค่าความถูกต้องและค่าความคลาดเคลื่อนกำลังสองเฉลี่ย MSE น้อยที่สุดคือ อัลกอริทึมโครงข่ายประสาทเทียมแบบ Multi-layer Perceptron (MLP) ให้ค่าความถูกต้องในการ Training Set สูงถึง 99.8125% และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย MSE เท่ากับ 0.00008649 และผลจากการ Training Set ของอัลกอริทึมโครงข่ายประสาทเทียม ให้ค่าความถูกต้องสูงถึง 99.5% และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย MSE เท่ากับ 0.00030976 ดังนั้นจึงสรุปได้ว่าอัลกอริทึมแบบ Multi-layer Perceptron (MLP) เหมาะกับข้อมูลชุดนี้มากที่สุด และสามารถนำผลที่ได้ไปพัฒนาต่อให้อยู่ในรูปแบบระบบสารสนเทศเพื่อช่วยวิเคราะห์โรคในสุกรได้อย่างมีประสิทธิภาพและแม่นยำมากขึ้น

5.2 ข้อเสนอแนะ

จากการทำงานของโครงข่ายประสาทเทียม จะเห็นได้ว่า โครงสร้างของระบบมีความสำคัญมากที่สุด เพราะถ้าออกแบบ โครงสร้างไม่สอดคล้องกับลักษณะของข้อมูลที่จะนำไปใช้งาน ก็จะพบว่าผลวิเคราะห์ออกมาจะไม่แม่นยำและมีความคลาดเคลื่อนสูง ดังนั้นจึงควรต้องคำนึงถึงโครงสร้างเป็นอันดับแรก โดยให้มีจำนวนนิวรอนและจำนวนชั้นที่เหมาะสมกับงาน จากนั้นจึงค่อยปรับค่าพารามิเตอร์ต่าง ๆ เพื่อเพิ่มประสิทธิภาพ ของการทำงานของโครงข่าย โดยในการพัฒนาครั้งต่อไป อาจจะเพิ่มตัวแปรอื่นด้วย เช่น ระดับความรุนแรงของอาการ เพื่อช่วยให้การจำแนกมีความถูกต้องแม่นยำยิ่งขึ้น

เอกสารอ้างอิง

- [1] จิตโสภิน บุญประเสริฐ และคณะ “สภาพการเลี้ยงดูของเกษตรกรที่ เข้าร่วม โครงการ สุก ร ชุน จ้าง เลี้ยง บริษัท เบทาโกร เกษตร อุดสาหกรรม สำนักงาน พิษณุโลก” การประชุมผลงานวิจัย ระดับบัณฑิตศึกษา มหาวิทยาลัยสุโขทัยธรรมมาธิราช ครั้งที่ 2, 2555
- [2] สุวรรณ พรหมทอง “การทำฟาร์มสุกร” สาขาเทคโนโลยีการเกษตร คณะวิทยาศาสตร์ มหาวิทยาลัยรามคำแหง, 2552
- [3] พบพร กระแสชัย “ระบบผู้เชี่ยวชาญแบบฟัซซี่เพื่อการวินิจฉัยโรคทาง การแพทย์” ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยเชียงใหม่, 2553
- [4] ภรณ์ชา อามฤรัตน์ และ พงษ์ มีสีง “การเปรียบเทียบประสิทธิภาพ การจำแนกกลุ่มข้อมูลของโรคลมร้อนด้วยวิธีการทางเครือข่ายประสาทเทียม” ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ และ ภาควิชาครุศาสตร์ไฟฟ้า คณะครุศาสตร์ อุดสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2553
- [5] จิตตภู พูลวัน. “ระบบวิเคราะห์โรคทั่วไปโดยใช้การจำแนกข้อมูลบนอินเทอร์เน็ต” ภาควิชาเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2550
- [6] สันติภาพ พันธุ์เสื่อ และมณเฑียร รัตนศิริวงศ์วุฒิ “ระบบการพยากรณ์ประเภทผู้ป่วยโรคเบาหวานโดยใช้โครงข่ายประสาทเทียม” สาขาวิชาระบบสารสนเทศเพื่อการจัดการ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2556
- [7] เพชรนรินทร์ แก้วหล้า “ระบบการพยากรณ์จำนวนผู้ป่วยโรคหัวใจโดยใช้โครงข่ายประสาทเทียม” ภาควิชาเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2553
- [8] เกียรติศักดิ์ จันทร์แก้ว และสุพจน์ นิตย์สุวรรณ. “การเปรียบเทียบผล การพยากรณ์อนุกรมเวลาราคาปาล์มน้ำมันโดยการใช้โครงข่ายประสาทเทียมฟังก์ชันพหุนามและโครงข่ายประสาทเทียม” สาขา ระบบสารสนเทศเพื่อการจัดการ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2555
- [9] นรินทร์ พนาवास และนิเวศ จิระวิชิตชัย. “การจำแนกมะเร็งเม็ดเลือดขาวโดยใช้โครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้น” สาขาวิชาคอมพิวเตอร์ธุรกิจ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยศรีปทุม วิทยาเขตชลบุรี, 2553

การเปรียบเทียบประสิทธิภาพวิธีการแบ่งกลุ่มข้อมูลที่ใช้สำหรับการวิเคราะห์กลุ่มอาหาร

The Comparison of Clustering Methods Used for Food Clustering Analysis

สุมาลี อีสริโยดม (Sumalee Eisariyodom)¹ และ อูรีรัฐ สุขสวัสดิ์ชน (Ureerat Suksawatchon)²

คณะวิทยาการสารสนเทศ มหาวิทยาลัยบูรพา

¹ann.eisariyodom@gmail.com, ²ureerat.w@gmail.com

บทคัดย่อ

งานวิจัยนี้ได้นำเสนอในลักษณะเชิงเปรียบเทียบเทคนิคที่ใช้ในการแบ่งกลุ่มเพื่อหาวิธีการที่เหมาะสมในการวิเคราะห์กลุ่มอาหาร ซึ่งประกอบด้วยวิธีการจัดกลุ่มแบบเคมีนส์ วิธีการจัดกลุ่มแบบฟัซซี่ซีมีนส์ และ วิธีการจัดกลุ่มแบบเอ็นโทรปีฟัซซี่ซีมีนส์ โดยมีการปรับปรุงการกำหนดจุดศูนย์กลางกลุ่มเริ่มต้นด้วยวิธีการแบบเคมีนส์ไลท์ โดยการวิเคราะห์กลุ่มอาหารในงานวิจัยนี้ได้ทำการวิเคราะห์จากสารอาหารหลักด้วยกัน 4 ชนิด ได้แก่ พลังงาน คาร์โบไฮเดรต ไขมัน และ โปรตีน โดยจะใช้เทคนิคการจัดกลุ่มกับข้อมูลอาหาร 5 ประเภท ได้แก่ อาหารจานหลัก อาหารว่าง ขนมหวาน ผลไม้ และเครื่องดื่ม รวมทั้งสิ้น 826 รายการอาหาร จากผลการทดลองพบว่าวิธีการแบบฟัซซี่ซีมีนส์ร่วมกับวิธีปรับปรุงจุดศูนย์กลางกลุ่มเริ่มต้นมีประสิทธิภาพการแบ่งกลุ่มข้อมูลที่ดีที่สุด เนื่องจากให้ค่า RS สูงที่สุดและค่า RMSSTD น้อยที่สุด

คำสำคัญ: การวิเคราะห์กลุ่มอาหาร, การแบ่งกลุ่ม, ฟัซซี่ซีมีนส์

Abstract

This research presents the comparison of clustering techniques to find the appropriate methods for food clustering analysis. We analyze with three clustering methods including K-means, Fuzzy C-means, and Entropy Fuzzy C-means algorithms. Before performing with these three methods, we start with cluster center initialization by using K-means-like algorithm. In this research, food clustering analysis is done with the four main nutrients which are energy, carbohydrate, fat, and protein in each category including main dishes, snacks,

desserts, fruits and drinks totally 826 food items. From the experimental results show that Fuzzy C-means with improving the cluster center initialization presents the best performance because it gives the highest RS and the minimum value of RMSSTD.

Keyword: Food Cluster Analysis, Clustering, Fuzzy C-means.

1. บทนำ

การเปลี่ยนแปลงพฤติกรรมมารับประทานอาหารของผู้บริโภคเป็นสาเหตุที่ทำให้ผู้บริโภคส่วนใหญ่มีภาวะโภชนาการเกิน รวมทั้งทำให้เสี่ยงต่อการเป็นโรคอ้วน โรคเบาหวาน ภาวะความดันโลหิตสูง [1] ซึ่งส่วนหนึ่งมาจากการที่รับประทานอาหารโดยไม่เลือก ไม่ได้คำนึงถึงว่าสิ่งที่รับประทานเข้าไปได้รับพลังงานเกินกว่าที่จะได้รับในแต่ละวันหรือไม่ อีกทั้งไม่ทราบว่าแต่ละวันนั้นร่างกายได้รับสารอาหารที่ครบถ้วนหรือไม่ ดังนั้นเพื่อให้ทราบว่าอาหารที่รับประทานนั้นสามารถรับประทานได้มากน้อยเพียงใด ผู้บริโภคจึงจำเป็นต้องจะต้องทราบว่า อาหารที่รับประทานนั้นจัดอยู่ในประเภทหรืออาหารกลุ่มใด ดังนั้นในงานวิจัยนี้จึงมีแนวคิดที่จะใช้การวิเคราะห์กลุ่มอาหารมาใช้ในการแบ่งอาหารออกเป็นกลุ่มตามลักษณะความเหมือนหรือคล้ายคลึงกันของสารอาหาร ซึ่งผลของการจัดกลุ่มอาหารทำให้ทราบว่าอาหารที่มีสารอาหารใกล้เคียงกันจะอยู่ในกลุ่มเดียวกัน

จากการศึกษาในงานวิจัยที่เกี่ยวข้องพบว่า มีหลายงานวิจัยที่นำเสนอเกี่ยวกับการวางแผนจัดมื้ออาหารและการจัดกลุ่มอาหาร เช่น งานวิจัยของ En-Yu และคณะ [2] ได้นำเสนอถึงการแนะนำอาหารที่จะช่วยให้นักโภชนาการให้คำปรึกษาแก่

ผู้บริโภคได้ดีขึ้น และช่วยให้ผู้บริโภคสามารถติดตามประวัติส่วนบุคคลของตนเองได้เอง โดยใช้กฎความสัมพันธ์ในการวางแผนรายการอาหารเพื่อให้ผู้บริโภคได้รับสารอาหารที่เหมาะสมและครบถ้วนตามหลักโภชนาการ นอกจากนี้ยังนำการวิเคราะห์ด้วยต้นไม้ตัดสินใจในการเลือกอาหารที่เหมาะสมให้กับผู้บริโภค

นอกจากนี้ยังมีงานวิจัยของ Yanisa และคณะ [3] ได้นำเสนอถึงการพัฒนาระบบให้คำปรึกษาด้านโภชนาการ สำหรับวางแผนรายการอาหาร โดยได้แบ่งอาหารออกเป็น 5 ประเภท ได้แก่ อาหารเช้า (Breakfast) อาหารจานหลัก (Main Dishes)¹ ขนมหวานและผลไม้ (Dessert and Fruits) ของว่าง (Snacks) และเครื่องดื่ม (Drinks) นอกจากนี้ยังนำวิธีการจัดกลุ่มแบบเคมีนส์ (K-means Clustering) มาใช้ในการแบ่งกลุ่มของอาหารที่มีลักษณะสารอาหารที่คล้ายกันจัดให้อยู่ในกลุ่มเดียวกัน และยังนำระบบผู้เชี่ยวชาญมาใช้ในการจัดกลุ่มอาหาร

จากงานวิจัย [3] ได้ใช้ขั้นตอนวิธีการจัดกลุ่มแบบเคมีนส์ แม้ว่าจะเป็นวิธีการที่นิยม เพราะเป็นวิธีการที่พัฒนาง่ายและไม่ซับซ้อน แต่ขั้นตอนวิธีการจัดกลุ่มแบบเคมีนส์นั้นยังมีข้อเสียอยู่หลายประการ คือ การหาจุดศูนย์กลางกลุ่มเริ่มต้นได้ใช้วิธีการสุ่มเลือก ซึ่งแต่ละครั้งจะได้จุดศูนย์กลางกลุ่มเริ่มต้นที่ไม่เหมือนกัน ทำให้การจัดกลุ่มอาหารไม่เหมือนกัน และไม่สามารถทราบได้ว่าการจัดกลุ่มในครั้งใดเหมาะสมที่จะนำมาใช้ในการวิเคราะห์ อีกทั้งอาจมีความคลาดเคลื่อนในการจัดกลุ่มอาหารออกมาได้ และในการเลือกกลุ่มก็ไม่ว่ากลุ่มไหนคือกลุ่มที่เหมาะสม

จากปัญหาที่เกิดจากงานวิจัยดังกล่าวข้างต้น งานวิจัยนี้จึงจะแก้ประเด็นปัญหาในการพัฒนาขั้นตอนการจัดกลุ่มอาหาร โดยงานวิจัยนี้เป็นการนำเสนอในลักษณะเชิงเปรียบเทียบเทคนิควิธีการแบ่งกลุ่มที่มีประสิทธิภาพกว่าวิธีการแบบเคมีนส์ โดยเริ่มต้นจะกำหนดจุดศูนย์กลางกลุ่มเริ่มต้นแต่ละกลุ่มด้วยวิธีการแบบเคมีนส์ไลต์ [4] เพื่อแก้ปัญหการกำหนดจุดศูนย์กลางกลุ่มเริ่มต้นที่เกิดจากขั้นตอนวิธีการแบบเคมีนส์แบบเดิม หลังจากนั้นจึงเป็นการจัดกลุ่มด้วยวิธีการดังต่อไปนี้ (1) วิธีการจัดกลุ่มแบบเคมีนส์ เพื่อเป็นการเปรียบเทียบประสิทธิภาพเมื่อมีการปรับปรุงวิธีการกำหนดจุดศูนย์กลางกลุ่มเริ่มต้น (2) วิธีการ

จัดกลุ่มแบบพีชชีมิมีนส์ [5] และ (3) วิธีการจัดกลุ่มแบบเอ็นโทรปีพีชชีมิมีนส์ [5] ซึ่งสองวิธีหลังนี้ใช้ความน่าจะเป็นในการกำหนดความเป็นสมาชิก (membership) ว่าข้อมูลแต่ละตัวควรจะอยู่กลุ่มใดมากกว่ากัน ซึ่งน่าจะให้ผลที่เหมาะสมกว่าการกำหนดความเป็นสมาชิกด้วยระยะทางเทียบกับจุดศูนย์กลางกลุ่ม ส่วนวิธีการเอ็นโทรปีพีชชีมิมีนส์ เป็นการนำฟังก์ชัน exponential มาปรับปรุงการหาฟังก์ชัน membership ของวิธีการพีชชีมิมีนส์ หลังจากที่เปรียบเทียบประสิทธิภาพของขั้นตอนวิธีการแบ่งกลุ่มทั้งสามวิธีที่ปรับปรุงวิธีการกำหนดจุดศูนย์กลางเริ่มต้น ก็จะนำวิธีการที่ดีที่สุดนำไปใช้ในการวิเคราะห์ข้อมูลของแต่ละกลุ่มอาหารต่อไป

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การวัดประสิทธิภาพการแบ่งกลุ่มข้อมูล [6]

- การวัดค่าความแตกต่างของข้อมูลที่เป็นอิสระต่อ

จำนวนของกลุ่ม (Davies-Bouldin หรือ DB)

DB เป็นการวัดประสิทธิภาพของการแบ่งข้อมูล ซึ่ง DB เป็นค่าสัดส่วนของผลรวมในกลุ่ม และระยะห่างระหว่างกลุ่ม ค่า DB ที่ดีที่สุดคือค่าที่น้อยที่สุด แสดงให้เห็นว่ามีการจัดกลุ่มที่ดีสามารถคำนวณค่า DB โดยใช้สมการดังนี้

$$DB = \frac{1}{n} \sum_{l=1}^n \max_{i \neq j} \left\{ \frac{S_n(Q_i) + S_n(Q_j)}{S(Q_i, Q_j)} \right\} \quad (1)$$

โดยที่

n คือ จำนวนกลุ่ม

$S_n(Q_i)$ คือ ค่าเฉลี่ยของระยะห่างระหว่างตัวแทนกลุ่มกับสมาชิกภายในกลุ่ม Q_i

$S_n(Q_j)$ คือ ค่าเฉลี่ยของระยะห่างระหว่างตัวแทนกลุ่มกับสมาชิกภายในกลุ่ม Q_j

$S(Q_i, Q_j)$ คือ ระยะห่างระหว่างตัวแทนกลุ่ม Q_i กับกลุ่ม Q_j

- การวัดค่าความแตกต่างของข้อมูลภายในกลุ่ม (Root Mean Square Standard Deviation หรือ RMSSTD)

การวัดค่าความแตกต่างของข้อมูลภายใน แสดงให้เห็นถึงประสิทธิภาพการจัดกลุ่มว่ามีมากน้อยเพียงใด ถ้าค่าความแตกต่างภายในกลุ่มน้อย หมายถึงการแบ่งกลุ่มที่ดี สามารถคำนวณค่าความแตกต่างภายในกลุ่ม โดยใช้สมการดังนี้

¹ อาหารจานหลัก (Main Dishes) คือ อาหารมื้อกลางวันและอาหารมื้อเย็น

$$RMSSTD = \sqrt{\frac{\sum_{i=1 \dots n_c} \sum_{k=1 \dots d}^{n_{ij}} (x_k - \bar{x}_j)^2}{\sum_{i=1 \dots n_c} (n_{ij} - 1)}} \quad (2)$$

โดยที่

n_c คือ จำนวนของกลุ่ม

n_{ij} คือ จำนวนของสมาชิกของแต่ละกลุ่ม

x_k คือ ข้อมูลของแต่ละกลุ่ม

\bar{x}_j คือ ค่าเฉลี่ยของจุดศูนย์กลางกลุ่มแต่ละสารอาหาร

- การวัดค่าความแตกต่างของข้อมูลระหว่างกลุ่ม

(R Squared หรือ RS)

การวัดค่าความแตกต่างของข้อมูลระหว่างกลุ่มนั้น เป็นอีกวิธีหนึ่งที่จะแสดงให้เห็นถึงประสิทธิภาพการจัดกลุ่มว่ามีมากน้อยเพียงใด ถ้าค่าความแตกต่างระหว่างกลุ่มมาก หมายถึงการแบ่งกลุ่มที่ดี โดยที่ค่าความแตกต่างระหว่างกลุ่มนี้จะอยู่ในช่วง 0 ถึง 1 สามารถคำนวณค่าความแตกต่างระหว่างกลุ่ม โดยใช้สมการดังนี้

$$RS = \frac{SS_t - SS_w}{SS_t} \quad (3)$$

เมื่อ $SS_t = \sum_{j=1}^d \sum_{k=1}^{n_j} (x_k - \bar{x}_j)^2 \quad (4)$

และ

$$SS_w = \sum_{i=1 \dots n_c} \sum_{k=1}^{n_{ij}} (x_k - \bar{x}_j)^2 \quad (5)$$

โดยที่

d คือ จำนวนสารอาหาร

n_j คือ จำนวนข้อมูลทั้งหมดของแต่ละสารอาหาร

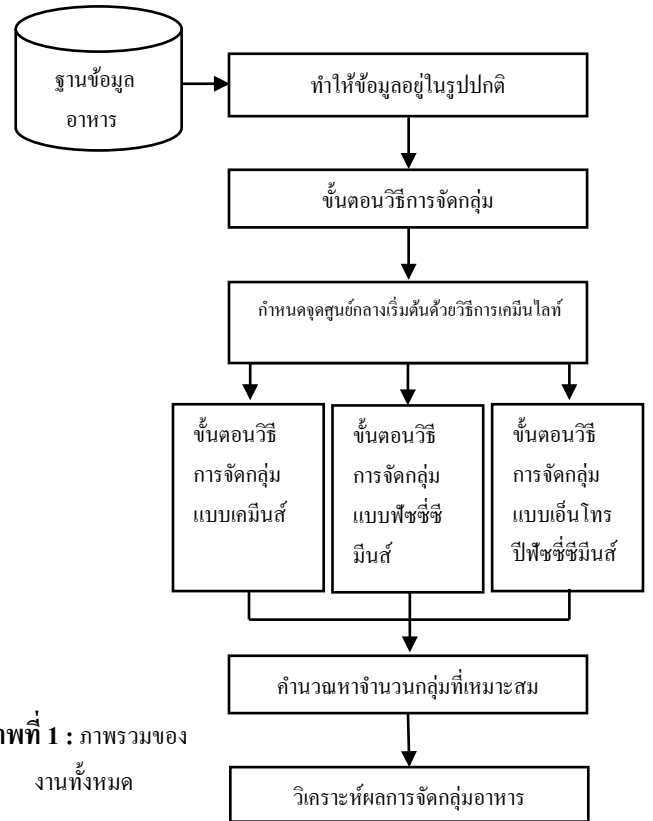
3. วิธีการที่นำเสนอในงานวิจัยนี้

ในงานวิจัยนี้เป็นการนำเทคนิคการจัดกลุ่มแบบเคมีนส์ไลต์ เทคนิคการจัดกลุ่มแบบพีชชีชีมินส์ และเทคนิคการจัดกลุ่มแบบเอ็นโทรปีพีชชีชีมินส์ที่มีการปรับปรุงการกำหนดจุดศูนย์กลางเริ่มต้น โดยภาพการทำงานโดยรวมเป็นดังภาพที่ 1

3.1 ลักษณะข้อมูลนำเข้า

การเก็บรวบรวมข้อมูลอาหารถูกรวบรวมจากโปรแกรม Inmucal Nutrients Program จากสถาบันวิจัยโภชนาการ มหาวิทยาลัยมหิดล ซึ่งได้มีการคำนวณสารอาหารแต่ละประเภทเอาไว้แล้ว สารอาหารที่นำมาใช้ ได้แก่ พลังงาน คาร์โบไฮเดรต ไขมัน และโปรตีน แบ่งเป็น 5 ประเภทได้แก่

อาหารจานหลัก อาหารว่าง ขนมหวาน ผลไม้ และเครื่องดื่ม รวมทั้งสิ้น 826 รายการอาหาร ดังตัวอย่างตารางที่ 1



ภาพที่ 1 : ภาพรวมของงานทั้งหมด

ตารางที่ 1 : ตัวอย่างข้อมูลอาหารจากโปรแกรม

รหัสอาหาร	พลังงาน	คาร์โบไฮเดรต	ไขมัน	โปรตีน
16001	259.1487	29.3724	7.2324	19.188
16002	677.8486	68.068	33.319	29.7154
16003	460.395	41.5125	23.3325	19.53
15001	114	11.08	7.3	1
15002	110.2	11.46	6.62	1.22
15003	109.8	10.68	6.9	1.24

3.2 การทำให้อยู่ในรูปมาตรฐาน (Data Normalization) [7]

การที่จะนำข้อมูลมาใช้ในการวิเคราะห์ต้องทำการจัดเตรียมข้อมูลโดยการนำข้อมูลมาทำให้อยู่ในรูปมาตรฐาน เพื่อปรับข้อมูลของตัวแปรต่างๆให้อยู่ในรูปแบบเดียวกัน ซึ่งการทำข้อมูลให้อยู่ในรูปมาตรฐานนี้จะมีค่าอยู่ในช่วง 0 ถึง 1 โดยมีกระบวนการ ดังนี้

- หาค่ามากที่สุด และค่าที่น้อยที่สุดของแต่ละสารอาหาร จากตารางที่ 1 ผลลัพธ์แสดงในตารางที่ 2

ตารางที่ 2 : ค่าที่มากที่สุดและค่าที่น้อยที่สุดของแต่ละสารอาหาร

	พลังงาน	คาร์โบไฮเดรต	ไขมัน	โปรตีน
X_{max}	906.2508	113.6	73.3	110.236
X_{min}	27.6369	0	0.28	0.8768

จากนั้นใช้สมการ (6) เพื่อปรับค่าของข้อมูล

$$x_{new} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (6)$$

โดยที่

x_{new} แทน ค่าที่ได้จากการปรับสเกล

x แทน ค่าที่ต้องการปรับสเกล

x_{max} แทน ค่าของข้อมูลที่ยิ่งใหญ่ที่สุดของสารอาหารนั้น

x_{min} แทน ค่าของข้อมูลที่ยิ่งน้อยที่สุดของสารอาหารนั้น

ผลของการปรับขอบเขตข้อมูลแสดงได้ดังตารางที่ 3

ตารางที่ 3 : ข้อมูลหลังการปรับสเกล

รหัสอาหาร	พลังงาน	คาร์โบไฮเดรต	ไขมัน	โปรตีน
16001	0.2635	0.2585599	0.095	0.167
16002	0.74004	0.5991901	0.452	0.264
16003	0.49255	0.3654269	0.316	0.171
15001	0.97896	0.7408027	0.978	0.036
15002	0.93898	0.7672798	0.883	0.049
15003	0.93477	0.712932	0.922	0.05

3.3 ขั้นตอนวิธีการจัดกลุ่ม

- การกำหนดจุดศูนย์กลางกลุ่มเริ่มต้นด้วยวิธีการเคมินไลท์ [8]

ในการจัดกลุ่มทั้ง 3 วิธีดังกล่าวมีข้อจำกัดอย่างหนึ่งของการกำหนดจุดศูนย์กลางกลุ่มเริ่มต้นของแต่ละกลุ่มตามจำนวนกลุ่มที่กำหนดไว้คือใช้วิธีการสุ่มจุดศูนย์กลางเริ่มต้นให้กับทุกกลุ่มซึ่งในบางครั้งจะมีศูนย์กลางกลุ่มเริ่มต้นบางตัวที่ได้จากการสุมนั้นอาจจะไปตกในบริเวณที่ห่างจากจุดศูนย์กลางกลุ่มตัวอื่นๆมาก ทำให้การแบ่งกลุ่มข้อมูลมีความผิดพลาดได้

เพื่อหลีกเลี่ยงปัญหาดังกล่าว งานวิจัยนี้จึงเลือกใช้ส่วนหนึ่งในขั้นตอนของเคมินไลท์ มาทำการกำหนดจุดศูนย์กลางกลุ่มเริ่มต้นโดยไม่ใช้วิธีการสุ่ม ด้วยการหาความคล้ายคลึงกันของ

ข้อมูลจากข้อมูลที่มีความใกล้ชิดกัน โดยหาจากการวัดระยะห่างระหว่างตัวแทนกลุ่มเริ่มต้น k กลุ่ม กับข้อมูล หากข้อมูลอยู่ใกล้กลุ่มใดมากที่สุดก็จะทำการเพิ่มให้เป็นสมาชิกในกลุ่มนั้น ขั้นตอนการหาจุดศูนย์กลางกลุ่มเริ่มต้น มีขั้นตอนดังนี้

ขั้นตอนที่ 1 กำหนดจำนวน k กลุ่ม

ขั้นตอนที่ 2 หาตัวแทนกลุ่มเริ่มต้น (Medoids) โดยการ

2.1 การหาระยะห่างแบบยูคลิด (Euclidean Distance) ระหว่างข้อมูลกับข้อมูล

2.2 คำนวณหา p_{ij} ดังสมการที่

$$P_{ij} = \frac{d_{ij}}{\sum_{i=1}^n d_{ij}} \quad (7)$$

โดยที่ P_{ij} คือ ค่าอัตราส่วนของระยะห่างระหว่างรายการอาหารที่ i กับรายการอาหารที่ j

d_{ij} คือ ค่าระยะห่างระหว่างรายการอาหารที่ i กับรายการอาหารที่ j

n คือ จำนวนรายการอาหารทั้งหมด

2.3 คำนวณค่าผลรวมของอัตราส่วนจากค่าระยะห่างของรายการอาหารแต่ละรายการ โดยใช้สมการ (15)

$$medoid_j = \sum_{i=1}^n P_{ij} \quad (8)$$

2.4 เรียงข้อมูลจากน้อยไปมาก แล้วนำค่ามาเป็นตัวแทนกลุ่มเริ่มต้นตามจำนวน k ที่ได้กำหนดไว้

เมื่อคำนวณค่า $medoid_j$ ของรายการอาหารแต่ละรายการครบแล้ว จากนั้นจึงจัดกลุ่มด้วยวิธีการทั้ง 3 วิธีแบบปกติต่อไป

- ขั้นตอนการจัดกลุ่มแบบเคมินส์ [9]

การจัดกลุ่มแบบเคมินส์เป็นเทคนิคการเรียนรู้แบบไม่มีผู้สอน (Unsupervised learning) ที่ง่ายที่สุดโดยจะแบ่งกลุ่มข้อมูลจากความคล้ายกันของข้อมูลและสามารถคำนวณได้จากการวัดระยะทางระหว่างเวกเตอร์ของข้อมูลโดยใช้การวัดระยะทางแบบยูคลิด (Euclidean Distance) โดยมีขั้นตอนดังนี้

ขั้นตอนที่ 1 กำหนดจำนวนกลุ่มและจุดศูนย์กลางเริ่มต้น โดยการสุ่มจุดศูนย์กลางเริ่มต้นขึ้นมา

ขั้นตอนที่ 2 คำนวณหาระยะทางโดยใช้ยูคลิดโดยใช้สมการดังนี้

$$Distance(p, q) = \sqrt{\sum_{k=1}^n (p_k - q_k)^2} \quad (9)$$

โดยที่

$Distance(p, q)$ คือ ระยะทางข้อมูล p กับ q

p_k คือ ค่าข้อมูล p ของข้อมูลที่ k

q_k คือ ค่าข้อมูล q ของข้อมูลที่ k

ขั้นตอนที่ 3 จัดกลุ่มให้กับข้อมูลและคำนวณจุดศูนย์กลางใหม่ โดยหาจากค่าเฉลี่ยข้อมูลทุกตัวที่อยู่ในกลุ่ม

ขั้นตอนที่ 4 ทำซ้ำขั้นตอนที่ 2 จนกว่าข้อมูลไม่เปลี่ยนกลุ่ม

• **ขั้นตอนการจัดกลุ่มแบบฟัซซีซีมินัส [5]**

การจัดกลุ่มแบบฟัซซีซีมินัสเป็นการยอมให้ข้อมูลในแต่ละกลุ่มมีการซ้อนทับกันหรือซ้ำกันได้ โดยอาศัยการให้ค่าการเป็นสมาชิกของข้อมูลต่อกลุ่มข้อมูลต่างๆ การจัดกลุ่มด้วยวิธีแบบฟัซซีซีมินัสเป็นเทคนิคในการจัดกลุ่มที่แก้ไขข้อเสียของการจัดกลุ่มด้วยวิธีแบบเคมินัส เนื่องจากการจัดกลุ่มวิธีแบบเคมินัสไม่เหมาะสมกับข้อมูลที่มีความสัมพันธ์กัน ซึ่งข้อมูลมีโอกาสเป็นสมาชิกเพียงกลุ่มใดกลุ่มหนึ่งเท่านั้น แต่การจัดกลุ่มแบบฟัซซีซีมินัสนั้นสมาชิกของกลุ่มมีโอกาสเป็นสมาชิกของข้อมูลระดับต่างๆ ในทุกๆกลุ่ม ซึ่งมีขั้นตอน ดังนี้

ขั้นตอนที่ 1 กำหนดกลุ่มข้อมูลที่ต้องการจัดกลุ่ม กำหนดค่าพารามิเตอร์ (m) ซึ่งต้องมากกว่าหนึ่ง และกำหนดจุดศูนย์กลางเริ่มต้นของข้อมูล

ขั้นตอนที่ 2 คำนวณค่าการเป็นสมาชิก (Membership) ของข้อมูลต่อกลุ่มข้อมูลต่างๆ โดยใช้สมการ ดังนี้

$$\mu_{ij} = \frac{\left(\frac{1}{d}\right)^{\frac{1}{m-1}}}{\sum_{i=1}^c \left(\frac{1}{d}\right)^{\frac{1}{m-1}}} \quad (10)$$

โดยที่

d คือ ระยะห่างระหว่างข้อมูลกับตัวแทนกลุ่มข้อมูล

c คือ จำนวนกลุ่มของข้อมูล

ซึ่งผลรวมของค่าการเป็นสมาชิกของข้อมูลสามารถคำนวณได้ดังสมการ

$$\sum_{i=1}^c \mu_{ij} = 1 \quad (11)$$

ขั้นตอนที่ 3 คำนวณจุดศูนย์กลางกลุ่มข้อมูลใหม่ โดยใช้สมการ ดังนี้

$$Z_i = \frac{\sum_{j=1}^n (\mu_{ij})^m x_j}{\sum_{j=1}^n (\mu_{ij})^m} \quad (12)$$

โดยที่ n คือ จำนวนข้อมูลทั้งหมด

ขั้นตอนที่ 4 ทำซ้ำขั้นตอนที่ 2 จนกว่าข้อมูลไม่เปลี่ยนกลุ่ม

• **ขั้นตอนการจัดกลุ่มแบบเอ็นโทรปีฟัซซีซีมินัส [5]**

การจัดกลุ่มด้วยวิธีแบบเอ็นโทรปีฟัซซีซีมินัส เป็นวิธีการหนึ่งที่ใช้ในการจัดกลุ่มข้อมูล ซึ่งมีวิธีการจัดกลุ่มข้อมูลที่

คล้ายคลึงกับการจัดกลุ่มแบบฟัซซีซีมินัส แต่เป็นการนำฟังก์ชัน exponential มาปรับปรุงการหาฟังก์ชัน membership ของวิธีการฟัซซีซีมินัส ซึ่งมีขั้นตอนดังนี้

ขั้นตอนที่ 1 กำหนดกลุ่มข้อมูลที่ต้องการจัดกลุ่ม กำหนดค่าพารามิเตอร์ (λ) ซึ่งต้องมากกว่าหนึ่ง และกำหนดจุดศูนย์กลางเริ่มต้นของข้อมูล

ขั้นตอนที่ 2 คำนวณค่าการเป็นสมาชิก (Membership) ของข้อมูลต่อกลุ่มข้อมูลต่างๆ โดยใช้สมการ ดังนี้

$$\mu_{ij} = \frac{\exp(-\lambda \times d_{ij})}{\sum_{u=1}^k (\exp(-\lambda \times d_{ij}))} \quad (13)$$

ซึ่งผลรวมของค่าการเป็นสมาชิกของข้อมูลสามารถคำนวณได้ดังสมการ

$$\sum_{i=1}^c \mu_{ij} = 1 \quad (14)$$

ขั้นตอนที่ 3 คำนวณจุดศูนย์กลางกลุ่มข้อมูลใหม่ โดยใช้สมการ ดังนี้

$$Z_i = \frac{\sum_{j=1}^n \mu_{ij} x_j}{\sum_{j=1}^n \mu_{ij}} \quad (15)$$

3.4 การกำหนดจำนวนกลุ่มที่เหมาะสม [10]

หลังจากที่ได้ทำขั้นตอนการจัดกลุ่มข้อมูลอาหารเรียบร้อยแล้ว งานวิจัยนี้ได้เลือกใช้การวัดค่า Silhouette Value มากำหนดจำนวนกลุ่มที่เหมาะสม โดยทดลองตั้งแต่จำนวนกลุ่มเท่ากับ 2 ถึง 9 กลุ่ม โดยใช้สมการ ดังนี้

$$S(i) = \frac{b(i)-a(i)}{\max\{a(i),b(i)\}}, \quad -1 \leq S(i) \leq 1 \quad (16)$$

โดยที่

$S(i)$ คือ ดัชนีความเชื่อมั่นกับค่าการเป็นสมาชิกของข้อมูลกับและจุดศูนย์กลางกลุ่มทั้งหมด

$a(i)$ คือ ระยะทางเฉลี่ยจากข้อมูลไปยังจุดศูนย์กลางกลุ่มทั้งหมด

$b(i)$ คือระยะทางเฉลี่ยที่น้อยที่สุดจากข้อมูลทั้งหมดไปยังจุดศูนย์กลางกลุ่มทั้งหมด

$$S_j = \frac{1}{m} \sum_{i=1}^m S(i) \quad (17)$$

โดยที่ m คือ จำนวนสมาชิกของแต่ละกลุ่ม

เมื่อได้ค่า S_j จะทำให้ได้ค่า Silhouette Value ดังสมการ

$$GS_u = \frac{1}{C} \sum_{j=1}^C S_j \quad (18)$$

โดยที่ C คือ จำนวนกลุ่ม

หลังจากทำการทดลองจำนวนกลุ่มที่มีค่า Silhouette Value มากที่สุด คือ กลุ่มที่เหมาะสมที่สุด ผลปรากฏว่า

วิธีการที่ 1) ขั้นตอนวิธีการแบบเคมีนส์ร่วมกับวิธีการแบบเคมีคอยส์ จำนวนกลุ่มที่เหมาะสม คือ 9 กลุ่ม มีค่า Silhouette Value เท่ากับ 0.7574

วิธีการที่ 2) ขั้นตอนวิธีการแบบฟัซซี่ซีมีนส์ร่วมกับวิธีการแบบเคมีคอยส์ จำนวนกลุ่มที่เหมาะสม คือ 7 กลุ่ม มีค่า Silhouette Value เท่ากับ 0.5563

วิธีการที่ 3) ขั้นตอนวิธีการแบบเอ็นโทรปีฟัซซี่ซีมีนส์ร่วมกับวิธีการแบบเคมีคอยส์ จำนวนกลุ่มที่เหมาะสม คือ 8 กลุ่ม มีค่า Silhouette Value เท่ากับ 0.6657

รองลงมาคือ วิธีการแบบเอ็นโทรปีฟัซซี่ซีมีนส์ร่วมกับวิธีการแบบเคมีคอยส์ และวิธีการแบบเคมีนส์ร่วมกับวิธีการแบบเคมีคอยส์มีประสิทธิภาพการแบ่งกลุ่มข้อมูลต่ำที่สุด

5. ขั้นตอนวิเคราะห์ผลการจัดกลุ่ม

จากหัวข้อที่ผ่านมาทำให้ทราบถึงวิธีการและจำนวนกลุ่มที่เหมาะสม คือ ขั้นตอนวิธีการแบบฟัซซี่ซีมีนส์ร่วมกับวิธีการแบบเคมีนส์ 7 กลุ่ม ซึ่งในการวิเคราะห์หว่าแต่ละกลุ่มอาหารมีลักษณะเด่นอะไรบ้าง ในงานวิจัยนี้ได้แบ่งระดับของสารอาหารออกเป็น 3 ช่วง ได้แก่ ระดับต่ำ ระดับปานกลาง และ ระดับสูง ดังตารางที่ 5

ตารางที่ 5 : การแบ่งช่วงของแต่ละสารอาหารทั้งหมด

	ต่ำ	ปานกลาง	สูง
พลังงาน	[27.6369-320.5082]	(320.5082-613.3795]	(613.3795-906.2508]
คาร์โบไฮเดรต	[0.0000-37.8667]	(37.8667-75.73333]	(75.7333-113.6000]
ไขมัน	[0.2800-24.6200]	(24.6200-48.9600]	(48.9600-73.3000]
โปรตีน	[0.8768-37.3299]	(37.3299-73.7829]	(73.7829-110.2360]

เมื่อทำการแบ่งช่วงของแต่ละสารอาหารแล้วทำการวิเคราะห์ออกมาแล้ว พบว่ามีบางกลุ่มที่มีระดับของสารอาหารซ้ำกันจึงทำการรวมกลุ่ม ดังตารางที่ 6

ตารางที่ 6 : กลุ่มอาหารจานหลักที่ทำการรวมกลุ่ม

	พลังงาน	คาร์โบไฮเดรต	ไขมัน	โปรตีน
กลุ่มที่ 1	ปานกลาง	ต่ำ	ต่ำ	ต่ำ
กลุ่มที่ 2	ต่ำ	ต่ำ	ต่ำ	ต่ำ
กลุ่มที่ 3	ต่ำ	สูง	ต่ำ	ต่ำ
กลุ่มที่ 4	ปานกลาง	สูง	สูง	สูง
กลุ่มที่ 5	สูง	ปานกลาง	ต่ำ	สูง

ภาพที่ 2 : ค่า Silhouette Value ในแต่ละวิธี

4. การวัดประสิทธิภาพการแบ่งกลุ่มข้อมูล

ในงานวิจัยนี้ได้นำการวัดประสิทธิภาพการแบ่งกลุ่มข้อมูลมาใช้เพื่อเปรียบเทียบเทคนิคการแบ่งกลุ่มทั้ง 3 วิธี ว่ามีการแบ่งกลุ่มที่ตีความน้อยเพียงใด ซึ่งได้เลือกใช้ 3 วิธี ได้แก่ 1) การวัดค่า DB 2) การวัดค่า RMSSTD และ 3) การวัดค่า RS

หลังจากทดลองแสดงการวัดประสิทธิภาพการแบ่งกลุ่มข้อมูล ดังตารางที่ 4

ตารางที่ 4 : ประสิทธิภาพการแบ่งกลุ่มข้อมูล

	DB	RMSSTD	RS
K-mean	0.4325982	0.59409458	0.90598641
Fuzzy C-mean	0.776179	0.48534274	0.92319604
Entropy	0.4548479	0.56335238	0.91085126

จากตารางที่ 4 แสดงให้เห็นว่าวิธีการแบบฟัซซี่ซีมีนส์ร่วมกับวิธีการแบบเคมีคอยส์มีประสิทธิภาพการแบ่งกลุ่มข้อมูลดีที่สุด เนื่องจากให้ค่า RS สูงที่สุด และ ค่า RMSSTD น้อยที่สุด

5.1 ผลการวิเคราะห์

ผลการวิเคราะห์ของงานวิจัยนี้ สามารถแบ่งกลุ่มอาหารได้ทั้งหมด 4 ประเภท ซึ่งจะยกตัวอย่างประเภทอาหารจานหลัก ดังนี้

อาหารจานหลัก สามารถแบ่งออกเป็น 5 กลุ่ม ได้แก่

กลุ่มที่ 1 : เป็นอาหารที่รับประทานได้แต่ต้องจำกัดปริมาณ เนื่องจากมีพลังงานค่อนข้างสูง ซึ่งอาจจะทำให้เกิดโรคต่างๆ ตามมา เช่น โรคอ้วน โรคเบาหวาน

กลุ่มที่ 2 : เป็นอาหารที่รับประทานได้ เนื่องจากเป็นกลุ่มที่สารอาหารต่างๆต่ำ ซึ่งผู้บริโภคสามารถรับประทานได้ปกติ

กลุ่มที่ 3 : เป็นอาหารที่รับประทานได้แต่ต้องจำกัดปริมาณ เนื่องจากมีคาร์โบไฮเดรตสูง ซึ่งมีผลทำให้ระดับน้ำตาลในเลือดเพิ่มขึ้น ทำให้เกิดโรคอ้วนและเสี่ยงต่อการเป็นโรคมากขึ้น

กลุ่มที่ 4 : เป็นอาหารที่ควรหลีกเลี่ยง หรือ บริโภคแต่น้อย เนื่องจากมีพลังงานค่อนข้างสูง คาร์โบไฮเดรตสูง ไขมันสูง ทำให้ระดับน้ำตาลในเลือดเพิ่มขึ้นได้ แต่มีโปรตีนสูง อาจจะช่วยรักษาสมดุลน้ำตาลในเลือดได้ เพราะกรดอะมิโนแอซิดในอาหารทำให้ปริมาณน้ำตาลในเลือดคงที่

กลุ่มที่ 5 : เป็นอาหารที่รับประทานได้แต่ต้องจำกัดปริมาณ หรือบริโภคแต่น้อย เนื่องจากมีพลังงานสูง คาร์โบไฮเดรตค่อนข้างสูง ทำให้ระดับน้ำตาลในเลือดเพิ่มขึ้นได้ แต่มีไขมันต่ำ และ โปรตีนสูง ช่วยรักษาสมดุลน้ำตาลในเลือดได้

6. สรุปผลการวิจัย

งานวิจัยนี้เป็นการหาวิธีที่เหมาะสมในการวิเคราะห์กลุ่มอาหาร โดยการเปรียบเทียบคัดเลือกจากเทคนิคการจัดกลุ่มข้อมูล 3 วิธี ได้แก่ ขั้นตอนวิธีการแบบเคมินส์ ขั้นตอนวิธีการแบบฟัซซี่ซีมินส์ และขั้นตอนวิธีการแบบเอ็นโทรปีฟัซซี่ซีมินส์ โดยทั้ง 3 วิธีได้มีปรับปรุงการกำหนดจุดศูนย์กลางกลุ่มเริ่มต้นก่อนการจัดกลุ่ม และใช้ค่า Silhouette Value ในการหาจำนวนกลุ่มที่เหมาะสมในแต่ละวิธี ส่วนในการคัดเลือกกว่าผลการจัดกลุ่มวิธีใดที่เหมาะสมในการนำมาใช้ในการวิเคราะห์กลุ่มอาหารต่อไปนั้น งานวิจัยนี้ได้ใช้ตัววัด ได้แก่ การวัด Davies-Bouldin การวัด RS และการวัด RMSSTD ซึ่งจากผลการทดลองพบว่าวิธีการแบบฟัซซี่ซีมินส์ร่วมกับวิธีการแบบเคมินส์มีประสิทธิภาพการแบ่งกลุ่มข้อมูลดีที่สุด เนื่องจากให้ค่า RS สูงที่สุดและค่า RMSSTD น้อยที่สุด ดังนั้นในงานวิจัยนี้จึงสรุปได้ว่าวิธีการแบบฟัซซี่ซีมินส์ที่มีการปรับปรุงจุดศูนย์กลางกลุ่มเริ่มต้นด้วยวิธีการแบบเคมินส์ไลท์ จึงเป็นวิธีการที่เหมาะสมที่จะนำมาใช้ในการวิเคราะห์ลักษณะของสารอาหารในแต่ละกลุ่มต่อไปและผลลัพธ์ของงานวิจัยจะนำไปใช้เพื่อ

พัฒนาระบบการวางแผนมื้ออาหาร (Menu Planning) ซึ่งจะนำกิจกรรมของแต่ละบุคคลมาเกี่ยวข้อง เพื่อนำมาพัฒนาในการแนะนำอาหารให้เฉพาะบุคคล

เอกสารอ้างอิง

- [1] ปาณบดี เอกะจัมปะกะ และ นิธิศ วัฒนมะโน, “พฤติกรรมกรรการบริโภคอาหารของประชาชนไทยในยุคโลกาภิวัตน์”, วารสารสถานการณ์สุขภาพไทย, สำนักนโยบายและยุทธศาสตร์ กระทรวงสาธารณสุข, 2552
- [2] L. En-Yu, Y. Don-Lin, and H. Ming-Chuan, “System Design of an Intelligent Nutrition Consultation and Recommendation Model,” *IEEE International Conference on Ubiquitous Intelligence and Computing and 9th International Conference on Autonomic and Trusted Computing*, 2012.
- [3] U. Yanisa, P. Suphakant, and C.Nagul, “Nutrition Counseling System and Food Menu Planning,” *Information Science and Applications (ICISA)*, 2010, pp.1-8.
- [4] H.S. Park, J.S. Lee and C.H. Jun, “A k-Means-Like Algorithm for k-Medoids Clustering and Its Performance”, *Department of Industrial and Management Engineering*, POSTECH, South Korea, Jun, 2009.
- [5] T. Kiatchai and J. Chuleerat, “Exponential Fuzzy C- Means for collaborative Filtering,” *Journal of Computer Science and Technology*, 2012.
- [6] K. Ferenc, L. Csaba, and B. Attila, “Cluster Validity Measurement Techniques,” *AIKED'06 Proceedings of the 5th WSEAS International Conference on Artificial Intelligence, Knowledge Engineering and Data Bases*, 2006, pp.388-393
- [7] วิระ จิริกิจอนุสรณ์ และบุตรีรัตน์ จรัสกุลชัย, “การลาดการณ์ ภาษีมูลค่าเพิ่มด้วยเทคนิคของเหมืองข้อมูล”, *วารสารเทคโนโลยีสารสนเทศ*, คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์, 2553
- [8] P. Hae-Sang, and J. Chi-Hyuck, “A Simple and fast algorithm for K-medoids clustering,” *Expert Systems with Applications* 36, 2009, pp 3336-3341
- [9] A. K. Jain, “Data Clustering: 50 years beyond K-means,” *Pattern Recognition Letters* 31, 2010, pp 651-666
- [10] N. Bolshakova and F. Azuaje, “Cluster validation techniques for genome expression data,” *Signal Processing* 83, 2013, pp.825-833

การวิเคราะห์หาความสัมพันธ์การจับคู่ในระบบพี่เลี้ยงทางไกลอิเล็กทรอนิกส์

The analysis of relationship-based clustering in e-Distance Mentoring System

พจนีย์ สุขชวานา (Photjane Sukchaona)¹และมนต์ชัย เทียนทอง (Monchai Teantong)²

^{1,2}สาขาวิชาคอมพิวเตอร์ศึกษา คณะคอมพิวเตอร์อุตสาหกรรมมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

Phot_it49@hotmail.com, monchai@kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อจัดกลุ่มนักศึกษาตามคุณลักษณะที่คล้ายกันและนำผลที่ได้ไปหาพี่เลี้ยงที่เหมาะสมในแต่ละกลุ่มโดยใช้เทคนิคการจัดกลุ่ม K-mean

ผลการวิจัยพบว่าการจัดกลุ่มตามคุณลักษณะที่คล้ายกันจัดกลุ่มได้ 4 กลุ่มดังนี้ กลุ่มที่ 1: เพศชาย, จบจากโรงเรียนการศึกษาระดับพื้นฐานและมีเกรดเฉลี่ยที่จบจากโรงเรียนเดิมมากกว่า 2.00 กลุ่มที่ 2: เพศหญิง, จบจากโรงเรียนการศึกษาระดับพื้นฐานและมีเกรดเฉลี่ยที่จบจากโรงเรียนเดิมมากกว่า 2.00 กลุ่มที่ 3: เพศหญิง, จบจากการศึกษานอกโรงเรียนและมีเกรดเฉลี่ยที่จบจากโรงเรียนเดิมมากกว่า 2.00 และกลุ่มที่ 4: เพศชาย จบจากโรงเรียนการศึกษาระดับพื้นฐานและมีเกรดเฉลี่ยที่จบจากโรงเรียนเดมน้อยกว่า 2.00 ผลที่ได้จากการจัดกลุ่มนำไปใช้จับคู่ระหว่างพี่เลี้ยงกับผู้เรียนตามความแตกต่างระหว่างบุคคลว่าส่งผลกระทบต่อการเรียนรู้ของผู้เรียนเพื่อใช้ในระบบพี่เลี้ยงทางไกลอิเล็กทรอนิกส์

คำสำคัญ: การจัดกลุ่มระบบพี่เลี้ยงทางไกลอิเล็กทรอนิกส์ การสังเคราะห์รูปแบบ

Abstract

The purpose of the study was to cluster similar characteristics of students and apply the results to find Mentor for each cluster by using K-means Technique.

In conclusion, Clustering Techniques were used to analyze similar characteristics of the students. There were 4 groups of participants. The results showed that the first group consisted of male, graduated from the School of Basic Education and the GPA to graduate from the same school more than 2.00, the second group consisted of female, graduated from the School of Basic

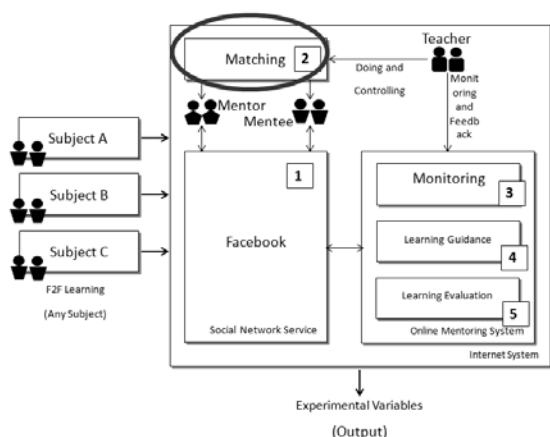
Education and the GPA to graduate from the same school more than 2.00, the third group consisted of female, graduated from the non-formal education and GPA. The graduates of the same school more than 2.00 and final group consisted of male, Graduate School of Basic Education and the GPA to graduate from the school not more than 2.00. The results of clustering applied to match mentor and students with differences characteristics of the students that affect the learning of the students to use in e-Distance Mentoring System.

Keyword: Clustering, e-Distance Mentoring System, Model synthesis.

1. บทนำ

การวิเคราะห์การจัดกลุ่มระหว่างพี่เลี้ยง (Mentor) กับผู้เรียน (Mentee) เพื่อหาความสัมพันธ์ในระบบพี่เลี้ยงทางไกลอิเล็กทรอนิกส์นี้เป็นส่วนหนึ่งของงานวิจัยเรื่องการสังเคราะห์โมเดลรูปแบบของระบบพี่เลี้ยงทางไกลอิเล็กทรอนิกส์โดยใช้บริการเครือข่ายสังคมออนไลน์ [1] ในส่วนประกอบ (Module) ที่ 2 ในการจับคู่ระหว่างพี่เลี้ยงกับผู้เรียน (Matching module) ของระบบพี่เลี้ยงทางไกล ดังภาพที่ 1 โดยมีส่วนประกอบทั้งหมด 5 โมดูลดังนี้ โมดูลที่ 1: ช่องทางการติดต่อสื่อสารผ่านทางเครือข่ายสังคมออนไลน์ (Social Network Service Module) โดยเลือกใช้เฟสบุ๊ก (Face book) เป็นช่องทางในการติดต่อสื่อสาร โมดูลที่ 2: การจับคู่ระหว่างพี่เลี้ยงกับผู้เรียน (Matching module) เป็นโมดูลที่นำเสนอในงานวิจัยนี้ โมดูลที่ 3: ส่วนที่ใช้สนับสนุนและส่งเสริม ติดตามการเรียนรู้ของผู้เรียนและพี่เลี้ยง และนำเสนอภารกิจงานที่มอบหมายให้แก่ผู้เรียน (Monitoring Module) โมดูลที่ 4: ส่วนแนะนำแนวทางการเรียนคอยช่วยเหลือ เก็บบันทึกกิจกรรมทางการเรียนและตรวจสอบประวัติการเข้าใช้ระบบ (Learning Guidance Module) และ

โมดูล 5: สำหรับประเมินผลกระบวนการเรียนรู้ของผู้เรียนที่เลี้ยงและผู้เรียน (Learning Evaluation Module) ซึ่งส่วนประกอบ 5 โมดูลข้างต้นได้มาจากวิธีการสนทนากลุ่ม (Focus group) จากผู้เชี่ยวชาญ จำนวน 10 คน หลังจากนั้นได้ให้ผู้เชี่ยวชาญอีกจำนวน 5 คน ประเมินความเหมาะสมของโมเดลรูปแบบที่เลี้ยงออนไลน์ทางไกลอิเล็กทรอนิกส์ ที่สังเคราะห์ขึ้น โดยใช้แบบสอบถาม ผลการประเมินความเหมาะสมรูปแบบระบบที่เลี้ยงทางไกลอิเล็กทรอนิกส์ จากผู้เชี่ยวชาญ ปรากฏว่า มีความเหมาะสมอยู่ในระดับมากที่สุด



ภาพที่ 1: กรอบแนวคิดการสังเคราะห์โมเดล

ผลลัพธ์ที่ได้จากงานวิจัยนี้นำไปใช้เป็นเครื่องมือในการเรียนการสอนที่สนองต่อพระราชบัญญัติการศึกษาแห่งชาติ พ.ศ.2542 และการปฏิรูปการศึกษาในทศวรรษที่สอง (พ.ศ. 2552 – 2561) ของการศึกษาไทยตามสภาพการเรียนรู้ของความแตกต่างระหว่างบุคคลและใช้เพื่อยืนยันผลการจับคู่ระหว่างที่เลี้ยงกับผู้เรียนว่าความแตกต่างระหว่างบุคคลส่งผลต่อการเรียนรู้ของผู้เรียน ซึ่งในส่วนของโมดูลที่ 2 ซึ่งเป็นส่วนระบบที่บริหารจัดการโดยผู้สอน เพื่อทำหน้าที่จับคู่ระหว่างที่เลี้ยงและผู้เรียนตามการจัดกลุ่มด้วยเทคนิค K-mean ในการจัดกลุ่มผู้เรียนกลุ่มละไม่เกิน 3 คนต่อที่เลี้ยง 1 คนตามเงื่อนไขของคุณสมบัติของที่เลี้ยงคือเป็นนักศึกษาชั้นปีที่ 2 หรือชั้นปีที่ 3 ที่เคยลงทะเบียนเรียนในรายวิชาที่ใช้ทดสอบในงานวิจัยนี้และได้เกรดตั้งแต่ B+ ขึ้นไป, มีความรู้ความสามารถในการใช้โปรแกรมด้านเครือข่ายสังคมออนไลน์ เช่น เฟสบุ๊ก, คุณสมบัติของผู้เรียนคือเป็นนักศึกษาชั้นปีที่ 1 คือต้องลงทะเบียนเรียนในรายวิชาเดียวกันกับรายวิชาที่ที่เลี้ยงเคยศึกษาแล้วและเงื่อนไข

การจัดกลุ่มคือเลือกรายวิชาจาก 3 ศาสตร์ที่แตกต่างกันนำมาใช้ทดสอบและใช้คุณสมบัติของนักศึกษาที่ได้จากเทคนิค K-mean เพื่อใช้ในการจัดกลุ่มและศึกษาถึงคุณลักษณะที่สัมพันธ์กันในการจัดกลุ่มที่เลี้ยงกับผู้เรียน

2. วัตถุประสงค์งานวิจัย

เพื่อจัดกลุ่มนักศึกษาตามคุณลักษณะที่คล้ายกันและนำผลที่ได้ไปหาที่เลี้ยงที่เหมาะสมในแต่ละกลุ่มโดยใช้เทคนิคการจัดกลุ่ม K-mean

3. วรรณกรรมที่เกี่ยวข้อง

3.1 การจัดกลุ่ม (Clustering)

การจัดกลุ่มเป็นการเรียนรู้แบบไม่มีผู้สอน (Unsupervised) เป็นการจัดกลุ่มต้องอาศัยแอททริบิวต์ในตัวแปรที่ศึกษาที่มีลักษณะคล้ายกันจัดให้อยู่กลุ่มเดียวกันและที่มีความต่างกันให้อยู่กลุ่มต่างกัน และในการวัดความคล้ายกันของแอททริบิวต์โดยทั่วไปพิจารณาจากระยะห่างระหว่างวัตถุ ซึ่งวิธีที่ได้รับความนิยมใช้กันคือการ วัดระยะห่างเชิงยูคลิดยกกำลังสอง (Squared Euclidean distance) ได้แก่ เทคนิค K-mean

การจัดกลุ่มแบบ K-mean เป็นวิธีการแบ่งข้อมูลออกเป็นกลุ่มย่อยตามจำนวนกลุ่ม (k) ที่กำหนดมีการคำนวณเพื่อหาระยะห่างระหว่างข้อมูลกับค่ากลางที่กำหนดหรือสุ่มขึ้นมาในแต่ละรอบจะจำแนกข้อมูลไปอยู่ในกลุ่มใดๆ ให้เลือกจากกลุ่มที่ข้อมูลมีระยะห่างจากค่ากลางน้อยที่สุด แล้วคำนวณค่ากลางของกลุ่มใหม่ กระทำซ้ำจนกระทั่งค่ากลางไม่เปลี่ยนแปลง ถ้ามีระยะห่างน้อยก็จัดข้อมูลนั้นเข้ากลุ่ม

การจัดกลุ่มอาศัยความเหมือน (Similarity) โดยใช้การวัดระยะห่างระหว่างจุดข้อมูลซึ่งมีวิธีที่นิยมใช้ได้แก่ การวัดระยะห่างแบบยูคลิด (Euclidean distance) และการวัดระยะห่างแบบแมนฮัตตัน (Manhattan distance)

การวัดระยะห่างแบบยูคลิดเป็นการหาระยะทางที่ใกล้ที่สุดระหว่างจุด p และ q โดย $p=(p_1, p_2, \dots, p_n)$, $q=(q_1, q_2, \dots, q_n)$ และ d เป็นระยะทางระหว่างจุด p, q

$$d(p, q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (1)$$

การวัดระยะห่างแบบแมนฮัตตันเป็นการหาระยะตามการลากเส้นตรงตามแนวแกนตั้งจากสองมิติระหว่างตำแหน่งสองตำแหน่ง

ค่าที่ใช้ในการเปรียบเทียบค่าความถูกต้องคือค่าผลรวมกำลังสองของความคลาดเคลื่อน (Sum of Square Error: SSE) เป็นค่าผลรวมของความคลาดเคลื่อนยกกำลังสองที่ควรมีค่าที่ต่ำที่สุดดังสมการที่ 2 เมื่อ x, m_i เป็นจุดข้อมูลในกลุ่ม c_i

$$SSE = \sum_{i=1}^K \sum_{x \in c_i} \text{dist}^2(m_i, x) \quad (2)$$

3.2 โปรแกรมด้านดาต้าไมน์นิง

โปรแกรมเวก้า (Waikato Environment for Knowledge Analysis: WEKA) เป็นโปรแกรมประเภทฟรีแวร์ (Freeware) สามารถทำงานได้หลายระบบปฏิบัติการ เป็นโปรแกรมที่ทำงานด้านดาต้าไมน์นิงให้เลือกใช้ได้อย่างครบถ้วนและสามารถเพิ่มฟังก์ชันเข้าในโปรแกรมได้ในภายหลังซึ่งในโปรแกรมนี้มีเทคนิค Cascade Simple KMeans เป็นวิธีที่เลือกหาค่าจำนวนกลุ่มที่ดีที่สุดโดยใช้เกณฑ์ Calinski and Harabasz criterion (CH criterion) โดยดูจากค่า Means CH score ที่มีค่าสูงเป็นค่าที่ดีที่สุดสำหรับการกำหนดจำนวนกลุ่มและมีค่าเริ่มต้นในการวัดระยะห่างแบบยุคลิด

โปรแกรม Orange canvas เป็นโปรแกรมด้านดาต้าไมน์นิงที่แสดงผลลัพธ์เป็นภาพกราฟิกที่สวยงาม สามารถทำงานได้ทุกระบบปฏิบัติการและจัดเป็นโปรแกรมฟรีแวร์ (Freeware)

งาน [2] นำเสนอการจัดกลุ่มข้อมูลในการจัดกลุ่มนักเรียนในการเรียนรู้ภาษาอังกฤษและภาษาฝรั่งเศส เพื่อให้สมาชิกในกลุ่มมีการช่วยเหลือซึ่งกันและกันในการเรียนรู้ภาษาและ [3] การจัดกลุ่มการเรียนรู้ไวยากรณ์ภาษา (grammar learning) ในเครือข่ายสังคมออนไลน์ดังเช่นเฟซบุ๊ก โดยการจัดกลุ่มใช้เทคนิค K-means เพื่อกำหนดคณกรณนักเรียนที่มีรูปแบบและความสามารถในระดับเดียวกันจากคุณลักษณะของนักเรียน เช่น อายุ ความรู้

โดย [4] ศึกษาเปรียบเทียบการคัดแยกกลุ่มข้อมูลชีวภาพของมนุษย์จากข้อมูลกิจกรรมส่วนบุคคลของคน 1 คนที่แสดงกิจกรรมแตกต่างกัน 11 พฤติกรรมที่จัดเก็บด้วยอุปกรณ์เซ็นเซอร์ โดยเปรียบเทียบระหว่างเทคนิค Simple K-means

และเทคนิค Two Step Cluster เพื่อใช้ในการคัดแยกกลุ่มข้อมูลพบว่าเทคนิค Simple K-means มีความสามารถในการแยกกลุ่มข้อมูลได้ดีมากและให้ค่าความถูกต้องสูงถึง 99.995%

จาก [5] ได้กำหนดขอบเขตและสร้างแบบประเมิณรูปแบบระบบที่เลี้ยงทางไกลอิเล็กทรอนิกส์ โดยใช้บริการเครือข่ายสังคมออนไลน์โดยผู้เชี่ยวชาญตรวจสอบ ดังนั้นก่อนนำแบบประเมิณไปทดลองใช้ผ่านบริการเครือข่ายสังคมออนไลน์นั้น ต้องมีการกำหนดที่เลี้ยงทางไกลอิเล็กทรอนิกส์ที่เหมาะสมกับกลุ่มผู้เรียน โดยศึกษาถึงคุณลักษณะในแต่ละกลุ่มให้เหมาะสมกับที่เลี้ยงทางไกลอิเล็กทรอนิกส์ โดยใช้บริการเครือข่ายสังคมออนไลน์ นำมาซึ่งประโยชน์ต่อทั้งผู้สอนและผู้เรียนโดยมีนิยามที่เกี่ยวข้องดังนี้

ที่เลี้ยง (Mentor) คืออาสาสมัครกลุ่มทดลองที่ใช้ระบบที่เลี้ยงทางไกล โดยใช้บริการเครือข่ายสังคมออนไลน์ ได้แก่ ผู้เรียนที่จบจากมหาวิทยาลัยราชภัฏกาญจนบุรีหรืออาสาสมัคร

นักศึกษา (Mentee) คือผู้เรียนกลุ่มทดลองที่ใช้ระบบที่เลี้ยงทางไกลอิเล็กทรอนิกส์ โดยใช้บริการเครือข่ายสังคมออนไลน์ ได้แก่ นักศึกษาระดับปริญญาตรี มหาวิทยาลัยราชภัฏกาญจนบุรี

จากงานวิจัยที่เกี่ยวข้องในงานวิจัยนี้ได้เลือกเทคนิค k-mean เพื่อใช้ในการจัดกลุ่มข้อมูลด้วยวิธีการหาจำนวนกลุ่มที่ดีที่สุดซึ่งในโปรแกรม Orange canvas ใช้เทคนิค k-means แบบ Optimized กับโปรแกรม WEKA ใช้เทคนิค cascade Simple KMeans เป็นอัลกอริทึมสำหรับการหา k ที่ดีที่สุดโดยงานวิจัยนี้ได้กำหนดช่วง k อยู่ระหว่าง 2 ถึง 10 และเปรียบเทียบการหาระยะห่าง Euclidean, Manhattan

4. วิธีการดำเนินการวิจัย

4.1 การเตรียมข้อมูล

ข้อมูลของมหาวิทยาลัยราชภัฏกาญจนบุรี ในชั้นปีที่ 1 และชั้นปีที่ 2 ภาคปกติ ระดับปริญญาตรี ปีการศึกษา 2555 ถึง 2556 จำนวน 2,192 ระเบียบในการประมวลผลข้อมูลประกอบด้วย แอททริบิวต์จำนวน 3 แอททริบิวต์ดังนี้

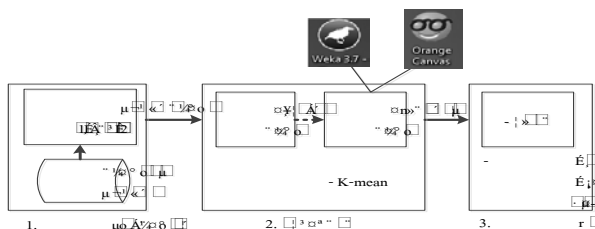
1) เพศ (sex) ประกอบด้วยค่า 1 หมายถึง 1 เพศชายค่า 2 หมายถึง เพศหญิง

2) ประเภทสถาบันที่จบการศึกษาในระดับอุดมศึกษาหรือเทียบเท่า (type_sc) ประกอบด้วยค่า 1 หมายถึงจบจากโรงเรียน การศึกษาขั้นพื้นฐาน (สพฐ.) ค่า 2 หมายถึงจบจากการศึกษา นอกโรงเรียน (กศน.) และค่า 3 หมายถึงจบจากวิทยาลัย เช่น วิทยาลัยเทคนิคกาญจนาบุรี

3) เกรดเฉลี่ยที่จบในระดับอุดมศึกษาหรือเทียบเท่า (avg_sc) ประกอบด้วยค่า low หมายถึงมีเกรดเฉลี่ยไม่เกิน 2.00 และ high มีค่าเกรดเฉลี่ยมากกว่า 2.00

4.2 กระบวนการดำเนินงาน

ขั้นตอนการทำงานแบ่งเป็น 3 ขั้นตอนดังนี้ 1 การรับข้อมูล เข้า 2) การประมวลผลและ 3) แสดงผลลัพธ์ดังภาพที่ 2



ภาพที่ 2: แสดงขั้นตอนการทำงาน

จากภาพที่ 2 ในขั้นตอนที่ 1 การรับข้อมูลเข้าเป็นขั้นตอน การนำข้อมูลจากฐานข้อมูลนักศึกษาเฉพาะชั้นปีที่ 1 และชั้นปีที่ 2 ภาคปกติ ระดับปริญญาตรี

ขั้นตอนที่ 2 การประมวลผลเป็นการศึกษาข้อมูลถึงความสัมพันธ์และเลือกของแอททริบิวท์ที่นำมาใช้ในการประมวลผล จากนั้นแปลงให้อยู่ในรูปแบบของโปรแกรม Orange และโปรแกรม WEKA เพื่อนำไปประมวลผลด้วยเทคนิค K-means ในขั้นตอนสุดท้ายเป็นการนำเอาผลลัพธ์ที่ได้มาวิเคราะห์และจัดกลุ่มเพื่อศึกษาถึงคุณลักษณะที่สัมพันธ์กันในการจัดกลุ่มที่เลี้ยง

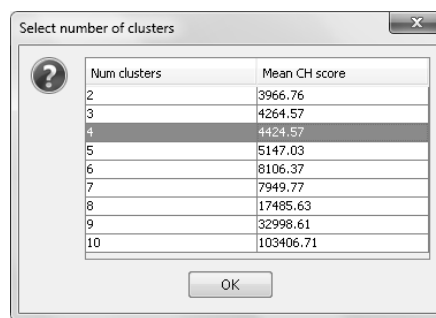
5. ผลการดำเนินงาน

ผลลัพธ์ของการประมวลผลข้อมูลด้วยเทคนิคการจัดกลุ่ม k-means โดยใช้โปรแกรม Orange canvas ที่มีเทคนิคการจัดกลุ่มแบบ k-means วิธี Optimal กล่าวคือสามารถกำหนดช่วงจำนวนกลุ่มเพื่อหาจำนวนกลุ่มที่ดีที่สุด แสดงผลลัพธ์ได้ดังตารางที่ 1 โดยกำหนดค่าเริ่มต้น(Initialization) เป็นแบบสุ่ม

ตารางที่ 1: การจัดกลุ่มด้วยเทคนิค K-mean โดยใช้ Orange canvas

Distance measures	k (best)	Score(Sihouette)
Euclidean	4	0.285
Manhattan	4	0.45

ในโปรแกรมเวก้าเทคนิค Cascade Simple KMeans เพื่อหาจำนวนกลุ่มข้อมูลที่ดีที่สุด ดังนั้นในการทดสอบนี้ได้กำหนดค่า manually Select Num Clustersเท่ากับ True เพื่อแสดงค่า Mean CH score ในแต่ละกลุ่มให้ผู้ใช้เลือกดังภาพที่ 3



ภาพที่ 3: หน้าต่างแสดงจำนวนกลุ่มและค่า Mean CH score

จากภาพที่ 3 พบว่าจำนวนกลุ่มที่ 1-4 ให้ค่า Mean CH score ที่สูงในลำดับที่ 4 ซึ่งใน [3] ได้ใช้เทคนิค Simple K-means ในโปรแกรมเวก้าให้ค่าความถูกต้องสูงถึง 99.995% ดังนั้นได้ทดสอบประมวลผลด้วยเทคนิคดังกล่าวโดยทดสอบจำนวนกลุ่มที่ 2 - 4 ดังตารางที่ 2 และพบว่าจำนวนกลุ่มเท่ากับ 4 ให้ค่า SSE ที่ต่ำที่สุดนั่นหมายถึงมีค่าความคลาดเคลื่อนน้อยดังนั้นในงานวิจัยนี้จึงเลือกจำนวนกลุ่มเท่ากับ 4 ซึ่งให้ผลลัพธ์ที่สอดคล้องกับเทคนิค K-mean โดยใช้ Orange canvas

ตารางที่ 2: เปรียบเทียบผลลัพธ์ที่ได้จากเทคนิค Simple K-means

จำนวนกลุ่ม	SSE	Time of build model (seconds)
2	1589.0	0.03
3	1169.	0.02
4	667.0	0.02

ตารางที่ 4: ผลการจัดกลุ่มด้วยเทคนิค K-Mean ที่ k = 4

Attribute	Cluster 1	Cluster 2	Cluster 3	Cluster 4
sex	1	2	2	1
type_sc	1	1	2	1
avg_sc	high	high	high	low

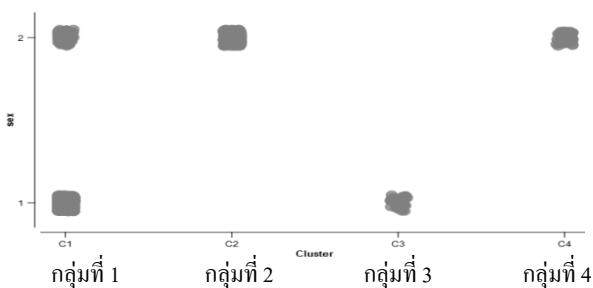
ผลการจัดกลุ่มนักศึกษาที่มีลักษณะคล้ายกันจัดได้ 4 กลุ่ม ได้แก่

กลุ่มที่ 1: เพศชาย, จบจากโรงเรียนการศึกษาขั้นพื้นฐาน และมีเกรดเฉลี่ยที่จบจากโรงเรียนเดิมมากกว่า 2.00

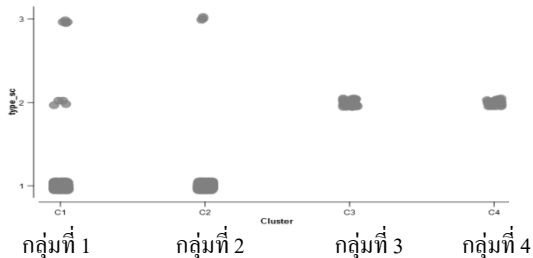
กลุ่มที่ 2: เพศหญิง, จบจากโรงเรียนการศึกษาขั้นพื้นฐาน และมีเกรดเฉลี่ยที่จบจากโรงเรียนเดิมมากกว่า 2.00

กลุ่มที่ 3: เพศหญิง, จบจากโรงเรียนการศึกษาขั้นพื้นฐาน และมีเกรดเฉลี่ยที่จบจากโรงเรียนเดิมมากกว่า 2.00

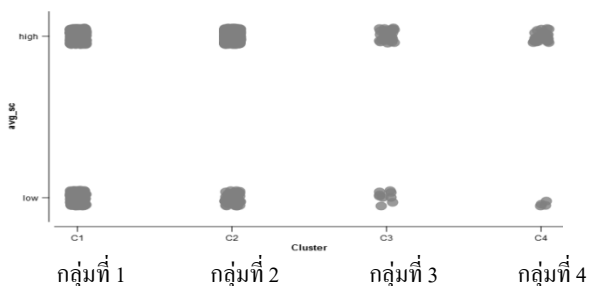
กลุ่มที่ 4: จบจากโรงเรียนการศึกษาขั้นพื้นฐานและมีเกรดเฉลี่ยที่จบจากโรงเรียนเดิมไม่เกิน 2.00



(ข) แสดงเพศในแต่ละกลุ่ม



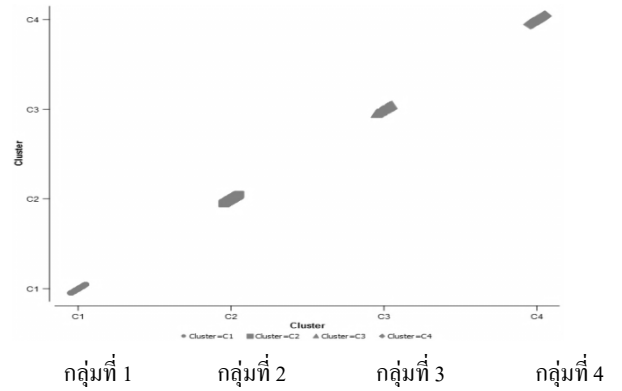
(ค) แสดงประเภทสถาบันที่จบการศึกษาหรือเทียบเท่าในระดับอุดมศึกษา หรือเทียบเท่าในแต่ละกลุ่ม



(ง) แสดงเกรดเฉลี่ยที่จบในระดับอุดมศึกษาหรือเทียบเท่าในแต่ละกลุ่ม

ภาพที่ 4: แสดงแอททริบิวต์ที่กระจายในแต่ละกลุ่ม

จากภาพที่ 4 แสดงแอททริบิวต์ที่กระจายในแต่ละกลุ่มข้อมูลกลุ่ม



ภาพที่ 5: แสดงกลุ่มข้อมูลในแต่ละกลุ่ม

จากภาพที่ 5 แสดงกลุ่มข้อมูลแต่ละกลุ่มที่มีขนาดกลุ่มใกล้เคียงกัน

6. สรุป

ในการวิจัยนี้ใช้เทคนิค k-means ในการจัดกลุ่มข้อมูล โดยใช้เทคนิค k-means วิธีแบบ Optimal เพื่อหาจำนวนกลุ่ม (k) ที่ดีที่สุดและใช้เทคนิค Cascade Simple KMeans ในโปรแกรมเวก้า เพื่อตรวจสอบผลลัพธ์จำนวนกลุ่มอีกครั้ง พบว่าจำนวนกลุ่มที่ดีที่สุดทั้งสองโปรแกรมคือจำนวนกลุ่มเท่ากับ 4 ดังนั้น ผลที่ได้จากการจัดกลุ่มจะนำไปใช้จับคู่ระหว่างที่เลี้ยงกับผู้เรียนตามความแตกต่างระหว่างบุคคลว่าส่งผลต่อการเรียนรู้ของผู้เรียนในส่วนของการ Matching Module ในงานวิจัยเรื่องการวิเคราะห์รูปแบบของระบบที่เลี้ยงทางไกลอิเล็กทรอนิกส์ โดยใช้บริการเครือข่ายสังคมออนไลน์ต่อไป

เอกสารอ้างอิง

- [1] พจนีย์สุขขานาและมนต์ชัยเทียนทอง, "การสังเคราะห์รูปแบบระบบที่เลี้ยงทางไกลอิเล็กทรอนิกส์ โดยใช้บริการเครือข่ายสังคมออนไลน์."
- [2] C. Troussas and M. Virvou, "Information Theoretic Clustering for an Intelligent Multilingual Tutoring System," *International Journal of Emerging Technologies in Learning*, vol. 8, pp. 55-61, 2013.
- [3] C. Troussas, et al., "Language Learning Assisted by Group Profiling in Social Networks," *International Journal of Emerging Technologies in Learning*, vol. 8, pp. 35-38, 2013.
- [4] อภิเชษฐ์บุญจาง, ชีรพงศ์เรือนน้อยและวนิดาแก่นอากาศ, "การศึกษาเปรียบเทียบเทคนิคเหมืองข้อมูลการคัดแยกกิจกรรมส่วนบุคคล,"

presented at the The 1st Nakhon Phanom University, 2011 Graduate Research Conference, 2554.

- [5] ปิยะรัตน์ คัญทัพ. "รูปแบบการสอนเพื่อพัฒนาทักษะด้านการคิดขั้นสูง โดยใช้กระบวนการเรียนการสอนแบบเว็บเควสในระดับประถมศึกษากรณีศึกษาโรงเรียนนานาชาติเกสินี" กรุงเทพฯ. วิทยานิพนธ์ปริญญาศึกษาศาสตรดุษฎีบัณฑิต สาขาวิชาหลักสูตรและการสอน มหาวิทยาลัยขอนแก่น, 2545

การสร้างภาพเอกสารไบนารีบนพื้นฐานของการทำคัสเตอร์ลำดับชั้นแบบโลคอล

A Document Binarization Method Based on Local Hierarchical Clustering

อมรเทพ แยมเกษร (Amorntape Yamkasorn)¹ และ อัมฉินุพันธ์ รอดทูกซ์ (Annupan Rodtook)²

ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยรามคำแหง

¹baracuda_tm@msn.com, ²sittisak168@yahoo.com

บทคัดย่อ

บทความวิจัยนี้นำเสนอวิธีการสร้างภาพไบนารีจากภาพเอกสารสีที่มีความซับซ้อนและความหลากหลายคุณลักษณะ วิธีการนี้ได้ประยุกต์แนวการดำเนินการแบบโลคอลด้วยการเชกเมนต์กลุ่มตัวอักษรในหน้าเอกสารและทำการแบ่งองค์ประกอบภายในเชกเมนต์ด้วยเทคนิคการทำคัสเตอร์แบบลำดับชั้น องค์ประกอบจะถูกวิเคราะห์และกำหนดคุณลักษณะการเป็นตัวอักษรหรือพื้นหลังในภาพไบนารี จากผลการทดลองพบว่าวิธีที่นำเสนอให้ค่าถูกต้องของผลลัพธ์สูงถึงร้อยละ 94.6 แสดงถึงประสิทธิภาพที่ดีกว่าเมื่อเทียบกับวิธีการอื่นๆ ที่เกี่ยวข้อง

คำสำคัญ: ภาพไบนารี การทำคัสเตอร์ลำดับชั้น ฟัซซี่ซีมีน

Abstract

This paper presents a method of document image binarization from color images consisting of complexity and a quality variety. This method applies a concept of local processing by segmenting character groups from a document page. Then, the components within a segment are detected using a method of hierarchical cluster technique. These components are analyzed and characterized as font or background in the binary image. The experimental results demonstrated that the proposed method provided 94.6% percentage high accuracy results which outperform to other related methods.

Keyword: Binary Image, Hierarchical Clustering, Fuzzy-C-Mean

1. บทนำ

ปัจจุบันนี้เทคโนโลยีสารสนเทศมีบทบาทความสำคัญในชีวิตประจำวันจนเรียกว่าเป็นยุคแห่งสารสนเทศ สารสนเทศ

เป็นเสมือนข้อมูลที่มีประสิทธิภาพเพื่อใช้สำหรับการตัดสินใจดำเนินงานในขั้นตอนสำคัญ สารสนเทศมาจากหลากหลายแหล่งข้อมูลและหนึ่งในนั้นคือ สื่อเอกสาร โดยสารสนเทศถูกบันทึกในรูปแบบที่หลากหลาย เช่น ตัวอักษรพิมพ์ ตัวอักษรลายมือเขียน ตัวเลข สัญลักษณ์ หรือรูปภาพ เป็นต้น

ในยุคโลกาภิวัตน์ที่มีความก้าวหน้าทางเทคโนโลยีและงานวิจัย สารสนเทศในสื่อสิ่งพิมพ์สามารถนำกลับมาใช้ใหม่หรือจัดเก็บในรูปแบบของแฟ้มข้อมูลเอกสารคอมพิวเตอร์ด้วยกรรมวิธีของระบบการรู้จำตัวอักษร (Optical Character Recognition, OCR) ปัจจัยสำคัญต่อประสิทธิภาพของระบบการจดจำรูปแบบคือ คุณภาพของเอกสาร กรณีเอกสารคุณภาพสูง การคัดแยกตัวอักษรออกจากพื้นหลังถูกดำเนินการอย่างมีประสิทธิภาพ และภาพผลลัพธ์มีความถูกต้องสูงคือ ภาพไบนารีที่ประกอบด้วยสององค์ประกอบคือ ตัวอักษร/ตัวเลข (จุดภาพสีดำ) และพื้นหลัง (จุดภาพสีขาว) ภาพดังกล่าวถูกใช้เป็นข้อมูลเริ่มต้นสำหรับระบบการรู้จำตัวอักษร

ปกติการประมวลผลภาพด้วยหลักการทำแธรชโวลด์ภาพ (Image Thresholding) ถูกนำมาใช้ในการสร้างภาพไบนารี ค่าแธรชโวลด์ถูกคำนวณหรือพิจารณาเลือกจากฮิสโตแกรมค่าระดับความเข้มของภาพ ค่าแธรชโวลด์จะแบ่งฮิสโตแกรมออกเป็น 2 ส่วนคือ ช่วงระดับความเข้มตัวอักษร และฉากหลังตามลำดับ ปัจจุบันการค้นคว้าและทำวิจัยเกี่ยวกับขั้นตอนการสร้างภาพไบนารีจากภาพเอกสารสียังเป็นประเด็นที่น่าสนใจ มีการเผยแพร่ผลงานวิจัยอย่างต่อเนื่อง โดยเฉพาะงานวิจัยในกรณีของภาพเอกสารคุณภาพต่ำ มีความซับซ้อนและยุ่งยากในองค์ประกอบพื้นหลัง เช่น รอยคราบ แสงเงา และความสมมาตรของสีตัวอักษรเนื่องจากปัญหาความส่องสว่าง ความถูกต้องของภาพเอกสารไบนารีส่งผลกระทบต่อประสิทธิภาพของระบบ OCR

การทบทวนวรรณกรรมวิจัยพบว่า C. Yu [1] นำเสนอวิธีการทำเรดโซลด์ภาพโดยประยุกต์อัลกอริธึม Otsu [1] กับฮิสโตแกรมค่าระดับความเข้มของภาพการ์เดียน ซึ่งผลลัพธ์ของหลักการนี้อยู่ในระดับที่น่าพอใจ คือ สิ่งไม่พึงประสงค์จะถูกกำจัดออกจากภาพไบนารี Er. Nirpjeet [2] นำเสนอบทความเชิงสำรวจวิธีการสร้างภาพเอกสารไบนารีด้วยอัลกอริธึมรูปแบบต่างๆ เช่น Otsu, Bernsen, Niblack, Yanowitz, Bruckstein และ Sauvola ซึ่ง Nirpjeet ทำการประเมินประสิทธิภาพของอัลกอริธึมเหล่านี้โดยการทดสอบกับชุดภาพเอกสาร และ K. Khurshid [3] นำเสนอวิธี Nick โดยทำการประยุกต์หลักการประมวลผลแบบโลคอลกับวิธีการของ Niblack ในรูปแบบของหน้าต่างการประมวลผลแบบเลื่อน (Sliding Window) สำหรับการทำเรดโซลด์ภาพเอกสาร และทำการประเมินประสิทธิภาพของ Nick ด้วยการเปรียบเทียบความถูกต้องของผลลัพธ์กับวิธีการอื่นๆ ที่เกี่ยวข้อง คือ Niblack, Sauvola, Wolf, และ Feng ตามลำดับ

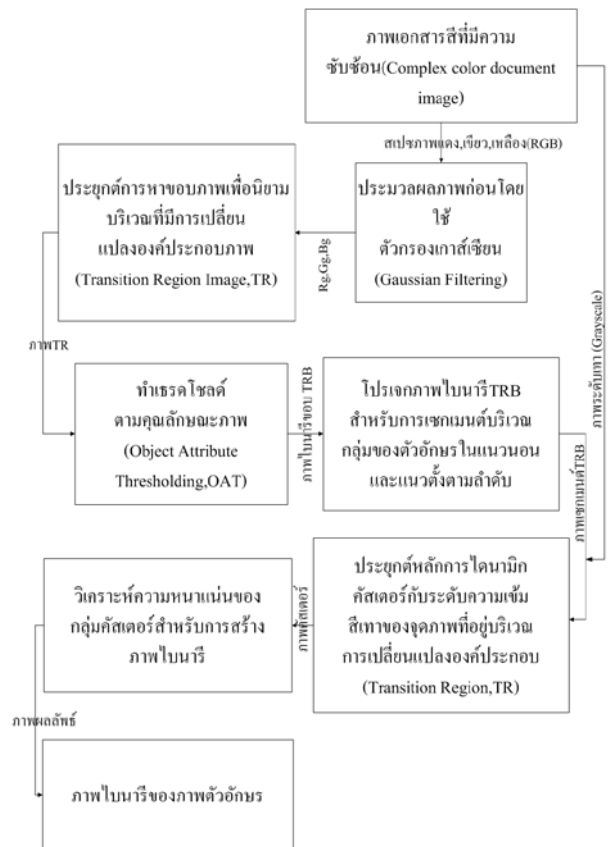
งานวิจัยในบทความนี้อยู่ในประเด็นปัญหาวิจัยเดิม คือ ภาพเอกสารสีคุณภาพต่ำ แต่เพิ่มประเด็นความซับซ้อนของเอกสาร เช่น กรณีความหลากหลายของสีตัวอักษรและพื้นหลัง กรณีระดับสีของพื้นหลังและตัวอักษรมีความใกล้เคียงกัน งานวิจัยนี้ประยุกต์หลักการงานแบบโลคอลร่วมการทำคัสเตอร์ข้อมูลแบบลำดับขั้น (Hierarchical Clustering) โดยภาพเอกสารถูกเชกเมนต์เป็นกลุ่มของตัวอักษร และทำคัสเตอร์ค่าระดับความเข้มของจุดภาพภายในเชกเมนต์ อัลกอริธึมฟัซซี่ซีมีน (Fuzzy C-Mean, FCM) ถูกนำมาประยุกต์ใช้ในการคัสเตอร์ข้อมูลเป็นลำดับขั้นในรูปแบบของเดนโดแกรม ผลลัพธ์ของการทำคัสเตอร์สะท้อนถึงกลุ่มข้อมูล (องค์ประกอบ) ที่เหมาะสมภายในเชกเมนต์ จำนวนองค์ประกอบมีความหลากหลายขึ้นกับจำนวนเลดสีตัวอักษรและพื้นหลัง ขั้นตอนสุดท้ายทำการวิเคราะห์และกำหนดคุณลักษณะให้กับกลุ่มข้อมูลในการสร้างภาพไบนารี ขั้นตอนการทำงานดังที่กล่าวข้างต้นอยู่บนพื้นฐานของการทำคัสเตอร์ลำดับขั้นแบบโลคอล (Local Hierarchical Clustering, LHC)

ภาพเอกสารสีที่มีคุณภาพหลากหลายและคุณลักษณะที่แตกต่างกันถูกนำมาใช้ในการทดสอบประสิทธิภาพของวิธี LHC โดยความถูกต้องของภาพผลลัพธ์ถูกพิจารณาจากเกณฑ์

การคำนวณ ค่าความถูกต้องเชิงบวก (TP) ความถูกต้องเชิงลบ (TN) ค่าความผิดพลาดเชิงบวก (FP) และค่าความผิดพลาดเชิงลบ (FN) ในรูปแบบที่กำหนด

2.การสร้างภาพไบนารีด้วยวิธี LHC

ภาพที่ 1 แสดงไดอะแกรมการสร้างภาพไบนารีจากภาพเอกสารสีด้วยวิธี LHC โดยภาพเอกสารถูกประมวลผลในสเปซสี RGB คือ $R(x, y), G(x, y), B(x, y)$ และค่าระดับเทา $f(x, y)$ ดังขั้นตอนต่อไปนี้



ภาพที่ 1: ไดอะแกรมการทำงานของอัลกอริธึม LHC

1. ประมวลผลภาพด้วยตัวกรองสัญญาณเกาส์เซียนดังสมการต่อไปนี้

$$\begin{aligned}
 R_g(x, y) &= g_\sigma(x, y) * R(x, y) \\
 G_g(x, y) &= g_\sigma(x, y) * G(x, y) \\
 B_g(x, y) &= g_\sigma(x, y) * B(x, y)
 \end{aligned} \tag{1}$$

โดย σ คือค่าความเบี่ยงเบนมาตรฐานของตัวกรองสัญญาณ

2. นิยามภาพการ์เดียน $TR(x, y)$ แสดงพื้นที่ขอบภาพ ขอบเป็นบริเวณที่มีการเปลี่ยนแปลงองค์ประกอบของภาพ

(Transition Region) โดย $TR(x, y)$ ถูกนิยามด้วยสมการ (2) และ (3) ดังนี้

$$\begin{aligned} Gr_R(x, y) &= |\nabla R_g(x, y)| \\ Gr_G(x, y) &= |\nabla G_g(x, y)| \\ Gr_B(x, y) &= |\nabla B_g(x, y)| \end{aligned} \quad (2)$$

∇ คือโอเปอเรเตอร์เกรเดียนต์

$$TR(x, y) = \max\{Gr_R(x, y), Gr_G(x, y), Gr_B(x, y)\} \quad (3)$$

3. การทำธรดโซลด์กับภาพ $TR(x, y)$

$$TRB(x, y) = \begin{cases} 255, & TR(x, y) > Thr \\ 0, & Otherwise \end{cases} \quad (4)$$

ค่าธรดโซลด์ Thr ที่ใช้ในสมการที่ 4 ถูกคำนวณด้วยวิธีการทำธรดโซลด์ตามคุณลักษณะภาพ (Object Attribute Thresholding, OAT) [4]

หลักการของ OAT คือ กระบวนการทำซ้ำของอัลกอริทึม Otsu บนฮิสโตแกรมเพื่อปรับค่าธรดโซลด์ที่เหมาะสมสำหรับแยกองค์ประกอบของภาพ $TR(x, y)$ คือ องค์ประกอบที่มีค่าเกรเดียนต์สูง (คือ ขอบตัวอักษร) และ องค์ประกอบที่มีค่าเกรเดียนต์ต่ำ (คือ พื้นหลังเอกสาร และขอบของสิ่งไม่พึงประสงค์)

ฮิสโตแกรมค่าระดับเข้มของภาพ $TR(x, y)$ ถูกแบ่งเป็น 2 กลุ่ม คือ C_0 (การเกรเดียนต์ต่ำ) และ C_1 (การเกรเดียนต์สูง) ด้วยค่าธรดโซลด์ T_1 ซึ่งคำนวณจากอัลกอริทึมของ Otsu และฮิสโตแกรมในช่วงค่า C_1 ถูกพิจารณาแบ่งต่อด้วยกรรมวิธีเดิมเป็นกระบวนการทำซ้ำจนกระทั่ง $T_n > A$ ดังนั้น $Thr = T_{n-1}$

A คือค่าเกรเดียนต์ (โดยประมาณ) ของขอบตัวอักษรโดย $A = \max\{TR(x, y)\} * \varphi$ (φ คือ ค่าสัมประสิทธิ์การถ่วงน้ำหนัก $\varphi \in [0, 1]$) ผลลัพธ์ของการทำธรดโซลด์แสดงดังภาพที่ 3 (ค)

4. การเชกเมนต์กลุ่มตัวอักษรในภาพ $TRB(x, y)$ ด้วยวิธีการโปรเจกชันฮิสโตแกรมของจุดภาพขอบในแนวนอน (แถว) และ แนวตั้ง (คอลัมน์) ตามลำดับ (ดังภาพที่ 3 (ง) และ (จ))

4.1 จุดภาพที่อยู่นอกกลุ่มเชกเมนต์ถูกพิจารณาเป็นจุดภาพพื้นหลัง (สีขาว) ในภาพไบนารี

4.2 เชกเมนต์ที่มีขนาดเล็กกว่าเกณฑ์ที่กำหนด (ขนาดตัวอักษรโดยประมาณ) เชกเมนต์ดังกล่าวถูกพิจารณาเป็นสิ่งไม่

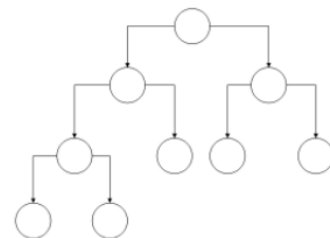
พึงประสงค์บนฉากหลัง ในกรณีที่ค่าระดับความเข้มใกล้เคียงกับระดับความเข้มของตัวอักษร จุดภาพภายในเชกเมนต์ดังกล่าวถูกกำหนดให้เป็นพื้นหลัง (สีขาว) ในภาพไบนารี

กลุ่มเชกเมนต์ที่ถูกกำหนดขึ้นภายในภาพ $TRB(x, y)$ จะถูกพิจารณาร่วมกับเชกเมนต์ในบริเวณเดียวกันของภาพเอกสารระดับเทา $f(x, y)$ ดังรายละเอียดในขั้นตอนต่อไป

5. ทำคัสเตอร์แบบลำดับขั้นกับข้อมูล คือ ค่าระดับเทา (Gray Level) ของจุดภาพภายในเชกเมนต์ กระบวนการในขั้นตอนนี้เป็นการประมวลผลร่วมระหว่างภาพ $f(x, y)$ และ $TRB(x, y)$ ดังนี้

5.1 พิจารณาค่าระดับเทาของจุดภาพเฉพาะบริเวณขอบภาพ คือ $f(x_i, y_i)$ ที่ $TRB(x_i, y_i) = 255$ โดย x_i, y_i เป็นพิกัดของจุดภาพบริเวณขอบ

5.2 ค่าระดับเทาจะถูกคัสเตอร์ออกเป็น 2 กลุ่มด้วยอัลกอริทึม ฟัซซีซีมีน (fuzzy-c-mean, FCM) [5] และดำเนินการต่อเนื่องเป็นลำดับขั้นจากบนลงล่าง (Top Down) ในรูปของเดนไดรแกรม [6] แสดงภาพที่ 2



ภาพที่ 2: เดนไดรแกรมของการคัสเตอร์

ค่า F-ratio [5] ถูกคำนวณเพื่อใช้พิจารณาการแบ่งกลุ่มข้อมูล (Splitting) โดยค่า

$$F_{ratio} = \sigma_B^2 / \sigma_w^2 \quad (5)$$

σ_B^2 คือ ค่าความแปรปรวนระหว่างกลุ่ม และ σ_w^2 ค่าความแปรปรวนภายในกลุ่ม การคัสเตอร์สองกลุ่มข้อมูลจะอยู่ในรูปแบบการทำซ้ำเป็นลำดับขั้น (ดังภาพที่ 2) และเกณฑ์ที่ใช้พิจารณายุติการคัสเตอร์ของแต่ละโหนดข้อมูล คือ $F_{ratio} < \alpha$ หลังจากเสร็จสิ้นขั้นตอนการแยกกลุ่มข้อมูล ทำการเรียงลำดับคัสเตอร์ตามค่าระดับเทาจากน้อยไปหามาก (พิจารณาค่ากลาง/ตัวแทน ของคัสเตอร์)

5.3 การรวมกลุ่มคัสเตอร์ (Merging) โดยพิจารณาจากคู่ของคัสเตอร์ในตำแหน่งติดกัน (Adjacent Clusters) ทำการประเมินระยะทางหรือความแตกต่างระหว่างคู่คัสเตอร์ และทำการรวมคัสเตอร์สำหรับคู่ที่มีความแตกต่างน้อยที่สุด ค่าระยะระหว่างคัสเตอร์ c_i และ c_j [6] ถูกคำนวณด้วยสมการดังนี้

$$Dist(c_i, c_j) = \sigma_B^2(c_i \cup c_j) \sigma_w^2(c_i \cup c_j) \quad (6)$$

ขั้นตอนการรวมคัสเตอร์อยู่ในรูปแบบของกระบวนการทำซ้ำ ทุกรอบการทำงานจำนวนคัสเตอร์จะลดลงหนึ่ง และมีการคำนวณค่ากลางสำหรับคัสเตอร์ที่เกิดจากการรวมกลุ่ม กระบวนการรวมกลุ่มจะยุติเมื่อ $Dist(c_i, c_j)$ ที่น้อยที่สุดมีค่ามากกว่าเกณฑ์ที่กำหนด จำนวนคัสเตอร์ที่ได้จากกระบวนการนี้แสดงถึงจำนวนกลุ่มข้อมูล (องค์ประกอบ) ที่เหมาะสมภายในเซกเมนต์ และใช้ค่ากลางของคัสเตอร์แทนค่าระดับเทาของกลุ่ม

5.4 พิจารณาทุกจุดภาพภายในเซกเมนต์ และ กำหนดกลุ่มให้จุดภาพดังกล่าว ด้วยการพิจารณาค่าความแตกต่างที่น้อยที่สุดระหว่างจุดภาพกับกลุ่มข้อมูล

$$d_i(x, y) = |f(x, y) - \rho_i|, i = 1, \dots, N \quad (7)$$

ρ_i คือ ค่าระดับเทาของกลุ่ม (องค์ประกอบ) ที่ i , N คือ จำนวนกลุ่ม ผลลัพธ์แสดงดังภาพที่ 5 (ฉ)

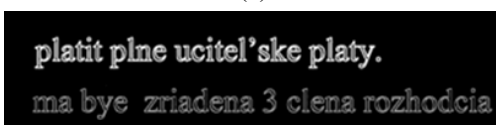
6. วิเคราะห์ความหนาแน่นของกลุ่มข้อมูลภายในเซกเมนต์ ด้วยการพิจารณาค่าความถี่ภายในกลุ่ม จุดภาพที่อยู่ในกลุ่มที่มีความถี่สูงสุดถูกพิจารณาเป็นจุดภาพพื้นหลัง (สีขาว) และจุดภาพในองค์ประกอบอื่นถูกพิจารณาเป็นจุดภาพตัวอักษร (สีดำ) โดยภาพผลลัพธ์แสดงดังภาพที่ 3 (ข)



(ก)



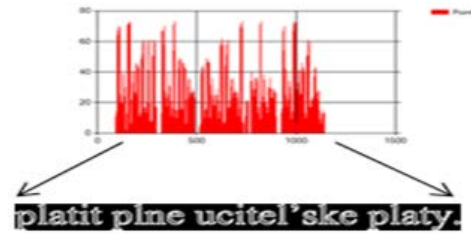
(ข)



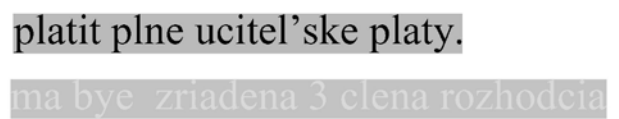
(ค)



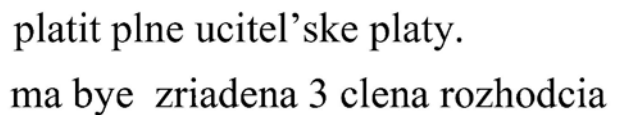
(ง)



(จ)



(ฉ)



(ช)

ภาพที่ 3: ผลการทำงานของวิธี LHC

- (ก). ภาพเอกสารสี (ข). ภาพ $TR(x, y)$ (ค). ภาพ $TRB(x, y)$
- (ง)-(จ). การโปรเจกชันฮิสโตแกรม (ง). แนวนอน (จ). แนวตั้ง
- (ฉ). องค์ประกอบภายในเซกเมนต์ (ช). ภาพไบนารี

3. ผลการดำเนินงานด้วยวิธี LHC

ทำการทดสอบวิธี LHC กับเซตของภาพเอกสารสี คือ IMG1 IMG2 IMG3 และ IMG4 ตามลำดับ (ดังภาพที่ 4 (ก)-(ง)) คุณลักษณะพิเศษของเอกสารชุดนี้ คือ ความหลากหลายของสีตัวอักษร สีของตัวอักษรและพื้นหลังมีความใกล้เคียงกันสูง สำหรับพารามิเตอร์ของอัลกอริทึมต่างๆ ถูกกำหนดดังนี้ ค่า σ ของตัวกรองสัญญาณเกาส์เซียนเท่ากับ 1.0 และค่า φ สำหรับการถ่วงน้ำหนักในขั้นตอน OAT เท่ากับ 0.75 ความถูกต้องของภาพผลลัพธ์จะถูกพิจารณาจากเกณฑ์การคำนวณต่อไปนี้ [7]

1. ค่าสัดส่วนความถูกต้องเชิงบวก (Sensitivity, ST)
2. ค่าสัดส่วนความถูกต้องเชิงลบ (Specificity, SPT)
3. ค่าคาดการณ์เชิงบวก (Positive predictive value, PPV)
4. ค่าความถูกต้องแม่นยำ (Accuracy, ACC)

$$ST = \left(\frac{TP}{TP + FN} \right) * 100 \quad (8)$$

$$SPT = \left(\frac{TN}{TN + FP} \right) * 100 \quad (9)$$

$$PPV = \left(\frac{TP}{TP + FP} \right) * 100 \quad (10)$$

$$ACC = \left(\frac{TP + TN}{TP + FP + FN + TN} \right) * 100 \quad (11)$$

ผลการคำนวณค่า *ST*, *SPT*, *PPV* และ *ACC* ถูกแสดงดังตารางที่ 1 ซึ่งค่าอัตราส่วนที่สูงจะเป็นดัชนีบ่งชี้ถึงความถูกต้องของผลลัพธ์ และสะท้อนถึงประสิทธิภาพของวิธีดังกล่าว การคำนวณข้างต้นถูกดำเนินการในขั้นตอนการเปรียบเทียบภาพผลลัพธ์กับภาพใบนารีอ้างอิง (Ground Truth Image) ค่าร้อยละของอัตราส่วนถูกนำไปใช้ในการเปรียบเทียบประสิทธิภาพของวิธี LHC กับวิธีการอื่นๆ ที่เกี่ยวข้องเช่น Bernsen, Nick, Niblack, Otsu และ Sauvola [8]

ตารางที่ 1: แสดงค่าอัตราส่วนความถูกต้อง (ร้อยละ)

ภาพเอกสารสี		Bernsen	Nick	Niblack	Otsu	Sauvola	LHC
IMG1	ST	62.39	100	5.38	100	100	100
	SPT	77.63	73.51	36.29	72.73	62.77	83.36
	PPV	72.43	63.96	36.78	62.50	40.70	80.04
	ACC	84.05	81.98	36.40	81.25	70.35	90.01
IMG2	ST	66.36	99.98	4.93	100	100	100
	SPT	84.73	79.21	39.73	85.61	65.09	86.96
	PPV	82.72	73.75	27.86	83.19	46.37	85.01
	ACC	89.31	86.87	37.71	91.59	73.18	92.20
IMG3	ST	74.96	99.93	8.67	95.22	100	90.53
	SPT	81.63	62.69	43.00	69.61	54.72	99.10
	PPV	78.36	40.48	31.79	56.52	17.28	99.11
	ACC	87.28	70.24	41.63	78.05	58.64	98.80
IMG4	ST	99.99	100	6.75	100	100	100
	SPT	83.75	80.72	47.86	83.12	57.33	94.75
	PPV	80.60	76.12	15.09	79.69	25.58	94.46
	ACC	90.30	88.06	46.52	89.84	62.79	97.21

จากผลการทดลองสร้างภาพใบนารีจากภาพเอกสารสี IMG1 ด้วยวิธีการต่างๆ (ภาพที่ 5 (ก)-(จ)) พบว่าวิธี LHC ให้ภาพผลลัพธ์ที่มีความถูกต้องสูง พิจารณาจากค่าการคำนวณ *ST*, *SPT*, *PPV* และ *ACC* ที่มีค่าเท่ากับ 100, 83.36, 80.04 และ 90.01 ตามลำดับ (ตารางที่ 1) กรณีของวิธี Bernsen, Nick และ Otsu สามารถให้ผลลัพธ์อยู่ในช่วงเกณฑ์ที่ดีแต่มีบางส่วนของ

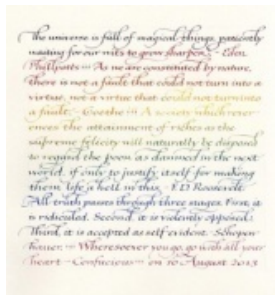
ตัวอักษรขาดหายไปใบบนารี ซึ่งเป็นผลกระทบจากความหลากหลายของสีตัวอักษร ค่าความถูกต้องของผลลัพธ์ (ACC) สำหรับวิธี Bernsen, Nick และ Otsu เท่ากับ 84.05, 81.98 และ 81.25 ตามลำดับ จากการพิจารณาผลการทดลองของภาพเอกสารทั้ง 4 ภาพ (IMG1 IMG2 IMG3 และ IMG4) พบว่าวิธี LHC สามารถให้ค่าร้อยละของ ACC สูงสุด (โดยเฉลี่ย) เท่ากับ 94.57 ถือเป็นตัวชี้วัดประสิทธิภาพของวิธีการที่นำเสนอ



(ก).



(ข).



(ค).



(ง).

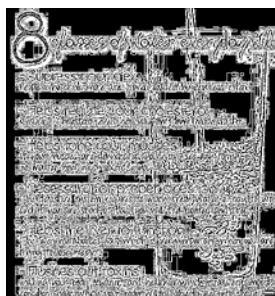
ภาพที่ 4: ภาพเอกสารสี (ก).IMG1(ข).IMG2 (ค).IMG3 (ง).IMG4



(จ).



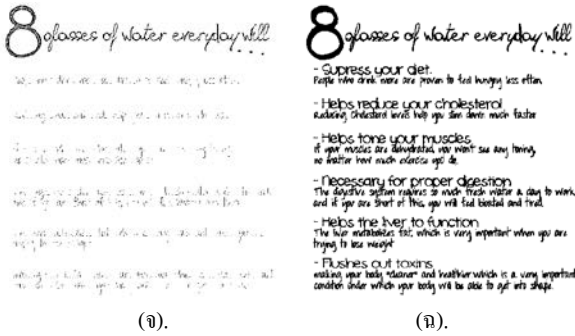
(ฉ).



(ช).



(ซ).



ภาพที่ 5: ภาพไบนารีของ IMG1

(ก). Bernsen (ข). Nick (ค). Niblack (ง). Otsu (จ).Sauvola (ฉ).LHC

4. สรุป

บทความวิจัยนี้เรานำเสนอวิธีสร้างภาพเอกสารไบนารีจากภาพเอกสารสีที่มีความหลากหลายคุณลักษณะ แนวคิดหลักของวิธีการนี้ คือ การทำงานแบบโลคอลโดยทำการเชกเมนต์กลุ่มตัวอักษรจากภาพ และทำการคัสเตอร์ค่าระดับความเข้มของจุดภาพแบบลำดับขั้นในแต่ละเชกเมนต์ จำนวนคัสเตอร์ที่เหมาะสมในแต่ละเชกเมนต์อาจไม่เท่ากันขึ้นอยู่กับระดับเจดสีตัวอักษร ขั้นตอนสุดท้ายทำการวิเคราะห์และกำหนดคุณลักษณะของแต่ละคัสเตอร์ว่าในรูปแบบของพื้นหลังหรือตัวอักษรในภาพผลลัพธ์ วิธีการที่นำเสนอถูกเปรียบเทียบกับวิธีการอื่นๆ ที่เกี่ยวข้อง ซึ่งจากผลการทดลองแสดงให้เห็นถึงประสิทธิภาพและความถูกต้องของภาพผลลัพธ์ที่ดีกว่า

เอกสารอ้างอิง

[1] C. Yu, and C. Lei, "Otsu's Thresholding Method Based on Gray Level-Gradient Two-Dimensional Histogram", *2010 2nd International Asia Conference on Informatics in Control*, Vol. 3, pp. 282-285, IEEE Press Piscataway, NJ, USA ©2010

[2] Er. Nirpjeetkaur, "A Review on Various Methods of Image Thresholding", *International Journal on Computer Science and Engineering (IJCSE)*, Vol. 3 No.10, pp. 3441-3443.

[3] K. Khurshid, I. Siddiqi, and N. Vincent "Comparison of Niblack Inspired Binarization Methods for Ancient Documents", *SPIE Proceeding Document Recognition and Retrieval XVI*, Vol. 7247, San Jose, CA, 18 January, 2009

[4] S. Chucherd and A. Rodtook, "Detection of Micro-calcifications in Mammograms Using the Object Attribute

Thresholding Algorithm", *Stamford Journal*, Vol. 4, No.1, January 2012 – May 2012, SSN 1906-1838.

[5] A. Rodtook and S.S. Makhanov, "Multi-Feature Gradient Vector Flow Snakes for Adaptive Segmentation of the Ultrasound Images of Breast Cancer", *Journal of Visual Communication and Image Representation*, Vol. 24, Issue 8, November 2013, pp. 1414-1430.

[6] A.Z. Arifin and A. Asano, "Image Segmentation by Histogram Thresholding using Hierarchical Cluster Analysis", *Pattern Recognition Letters*, Vol. 27, Issue 13, 1 October 2006, pp. 1515-1521

[7] A. Sopharak, B. Uyyanonvara, and S. Barman, "Automatic Exudate Detection for Diabetic Retinopathy Screening", *Science Asia Journal* Vol. 35, No. 1, March 2009, pp.80-88

[8] J. Motl, "Global and Local Image Thresholding Methods", [online] 2013, Available from: <http://www.mathworks.com/matlabcentral/fileexchange>

เครื่องมือแสดงภาพไทม์ไลน์สำหรับข่าวโดยอิงการจำแนกประเภทข่าว

News Timeline Visualization Tool Based on News Category Classification

วิทวัส สัจจเจริญพงษ์ (Vitthawat Sajjacharoenpong)¹ และ ทวิตตี้ เสนีวงศ์ ณ อยุธยา (Twittie Senivongse)²

ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

¹vitthawat.s@student.chula.ac.th, ²twittie.s@chula.ac.th

บทคัดย่อ

ในปัจจุบันมีการเผยแพร่ข้อมูลข่าวสารผ่านทางเครือข่ายอินเทอร์เน็ตเป็นจำนวนมาก หลายสำนักข่าวหรือแหล่งข่าวมีการจัดเตรียมบริการแบ่งปันข้อมูลข่าวสารผ่านเว็บในรูปแบบอาร์เอสเอสไว้ และต่างมีการจำแนกประเภทข้อมูลข่าวจำนวนมากนั้นออกเป็นหมวดหมู่ที่แตกต่างกัน อีกทั้งการนำเสนอข่าวส่วนใหญ่ยังอยู่ในรูปของรายการข่าวเรียงตามลำดับในแต่ละประเภทข่าว หรืออยู่ในรูปของกรอบข่าวแบ่งตามเนื้อเรื่องหรือประเภทข่าว การนำเสนอในลักษณะนี้ทำให้เห็นภาพรวมของข่าวที่หลากหลายในขณะหนึ่ง แต่หากต้องการติดตามข่าวเรื่องหนึ่งว่ามีพัฒนาการตามเวลาที่ผ่านไปอย่างไร จะทำได้ไม่สะดวกนัก งานวิจัยนี้เสนอทางเลือกในการติดตามข่าวโดยใช้เครื่องมือแสดงภาพไทม์ไลน์สำหรับข่าวโดยอิงการจำแนกประเภทข่าว เครื่องมือจะนำเข้าสู่ข้อมูลข่าวในรูปของอาร์เอสเอสจากหลายสำนักข่าว แล้วทำการจำแนกข่าวตามประเภทข่าว ด้วยวิธีการจำแนกข้อความแบบซีเอฟ-ไอดีเอฟแบบถ่วงน้ำหนัก การจำแนกจะอ้างอิงออนโทโลยีประเภทข่าวซึ่งผู้วิจัยสร้างขึ้นตามมาตรฐานโอพีทีซี ร่วมกับการเพิ่มคำศัพท์เสริมจากฐานข้อมูลเวิร์ดเน็ต เครื่องมือสามารถค้นหาข่าวตามประเภทที่จำแนกไว้และ/หรือตามคำสำคัญ และแสดงภาพไทม์ไลน์เพื่อแสดงพัฒนาการของข่าวเรื่องที่สนใจในเวลาที่ผ่านมาได้ การประเมินผลการจำแนกข่าวของเครื่องมือค่อนข้างน่าพอใจ โดยได้ค่าความแม่นยำ 72.52% ค่าเรียกคืน 66% และค่าเอฟ-เมเชอร์ 69.11%

คำสำคัญ: การแสดงภาพ การจำแนก ออนโทโลยี ข่าว

Abstract

At present, the Internet has been a major distribution channel for online news and information. Many news agencies or news sources provide news contents and

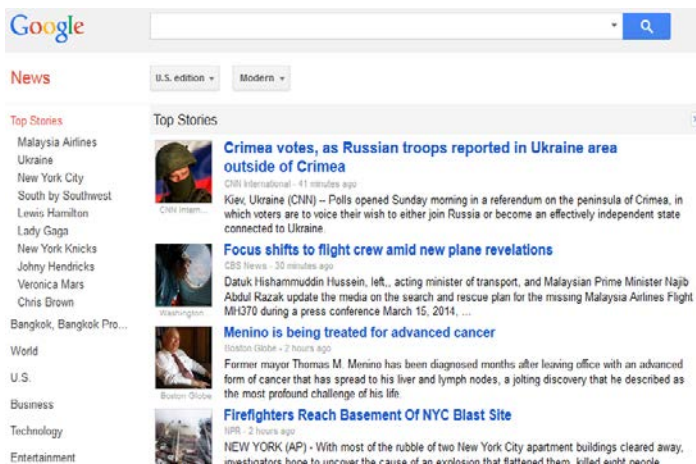
metadata for exchange via the Web in the RSS format, while using different ways to categorize this large amount of news. Moreover, online news is usually presented as a list of news items or as blocks of news under different categories or topics. Such presentation can give the overall picture of what is happening at the moment but it would be difficult for a news reader who needs to follow the development of a particular news topic over time. This research presents a news timeline visualization tool based on news category classification, as an alternative for online news readers. The tool takes RSS news feeds from different news agencies and classify them using the weighted CF-IDF text classification method. The classification is based on the news category ontology which we build to align with the IPTC standard and enhance with terms from the WordNet lexical database. The tool can search for news items by the category and/or keyword and visualize the news results in a timeline to show how the news develops over time. The evaluation of the classification is quite satisfactory having 72.52% precision, 66% recall, and 69.11% F-measure.

Keywords: Visualization, Classification, Ontology, News

1. บทนำ

ในแต่ละวันข้อมูลข่าวสารมีการพัฒนาและเผยแพร่ผ่านทางเครือข่ายอินเทอร์เน็ตเป็นจำนวนมาก แต่ละเว็บไซต์ข่าวได้มีการจัดเตรียมบริการการแบ่งปันข้อมูลข่าวระหว่างเว็บในรูปแบบอาร์เอสเอส (RSS) [1] เพื่อใช้สำหรับการเข้าถึงข้อมูลข่าวจากเว็บไซต์ข่าว โดยไม่จำเป็นต้องเข้าไปในเว็บไซต์ข่าวนั้น แต่ละเว็บไซต์ข่าวได้มีการจำแนกประเภทข้อมูลข่าว

จำนวนมากนั้นออกเป็นหมวดหมู่ที่แตกต่างกัน อีกทั้งการนำเสนอข่าวส่วนใหญ่ยังอยู่ในรูปของรายการข่าวเรียงตามลำดับในแต่ละประเภทข่าว หรืออยู่ในรูปของกรอบข่าวแบ่งตามเนื้อเรื่องหรือประเภทข่าว ภาพที่ 1 แสดงเว็บไซต์กูเกิลนิวส์ที่มีลักษณะเป็นพอร์ตัลรวบรวมข่าวจากหลายแหล่ง มีการจัดกลุ่มข่าวเรื่องเดียวกันจากหลายแหล่งไว้ด้วยกัน มีเมนูตามการจัดแบ่งหมวดหมู่ข่าวของกูเกิลนิวส์ และแสดงข่าวเป็นรายการเรียงลำดับ การนำเสนอในลักษณะนี้ทำให้เห็นภาพรวมของข่าวที่หลากหลายในขณะหนึ่ง แต่หากต้องการติดตามข่าวเรื่องหนึ่งว่ามีพัฒนาการตามเวลาที่ผ่านไปอย่างไร จะทำได้ไม่สะดวกนัก



ภาพที่ 1: ตัวอย่างประเภทข่าวและการนำเสนอข่าว

ไอพีทีซี (IPTC: International Press Telecommunications Council) ซึ่งเป็นกลุ่มของสำนักข่าวสำคัญของโลก ได้มีการพัฒนามาตรฐานสำหรับการแลกเปลี่ยนข่าวสารขึ้น [2] โดยแบ่งประเภทของข่าวตามลักษณะข่าวออกเป็น 17 ประเภท แต่ละประเภทมีคำศัพท์ที่เกี่ยวข้องระบุอยู่ด้วย งานวิจัยนี้เล็งเห็นความสำคัญในการรวบรวมข้อมูลข่าวสารจากเว็บไซต์สำนักข่าว จึงมีวัตถุประสงค์ที่จะพัฒนาเครื่องมือจำแนกประเภทข่าว เพื่อให้ผู้ใช้สามารถเข้าถึงข้อมูลข่าวตามประเภทข้อมูลข่าวจากหลายสำนักข่าวไปพร้อมกันได้ และสามารถติดตามพัฒนาการของข่าวเรื่องหนึ่ง ๆ ได้ ในการจำแนกประเภทนั้นจะใช้ประเภทข่าว 17 ประเภทของไอพีทีซีมาสร้างเป็นออนโทโลยีประเภทข่าว และใช้ฐานข้อมูลคำศัพท์ภาษาอังกฤษเวิร์ดเน็ต (WordNet) [3] มาเป็นส่วนเสริมในการสร้างออนโทโลยีให้สมบูรณ์ยิ่งขึ้น จากนั้นจึงใช้วิธีจำแนกข้อความแบบซีเอฟ-ไอดีเอฟแบบถ่วงน้ำหนัก (Weighted CF-IDF: Weighted Concept

Frequency - Inverse Document Frequency) [4] ซึ่งเป็นการจำแนกข่าวตามคำศัพท์หรือคอนเซปต์ในออนโทโลยีประเภทข่าว เครื่องมือสามารถค้นหาข่าวตามประเภทที่จำแนกไว้และ/หรือตามคำสำคัญที่ผู้ใช้สนใจ แล้วนำเสนอโดยการแสดงภาพ (Visualization) เพื่อการสื่อสารข้อความ โดยใช้เทคนิคไทม์ไลน์ (Timeline) [5] แสดงแถบข่าวที่มีวันหรือเวลาและข่าวที่เกิดขึ้นกำกับอยู่ ซึ่งจะทำให้ผู้ใช้สามารถทำความเข้าใจ โดยเฉพาะสามารถเห็นภาพหรือติดตามพัฒนาการของข่าวเรื่องที่สนใจในเวลาที่ผ่านมาได้สะดวกขึ้น ในส่วนท้ายของบทความจะแสดงการประเมินผลด้วยค่าความแม่นยำ ค่าระลอก ค่าเอฟ-เมเชอร์ และอภิปรายผล

2. งานวิจัยที่เกี่ยวข้อง

หลายงานวิจัยทำการจำแนกประเภทข่าวโดยใช้วิธีต่าง ๆ อาทิเช่น งานวิจัยของ Saha และคณะ [6] นำเสนอวิธีการจำแนกประเภทของอาร์เอสเอสฟีดจากเว็บไซต์สำนักข่าว โดยใช้แบบจำลองสำหรับสอน (Training Model) ในการจำแนกประเภทข่าว ก่อนให้บริการข้อมูลข่าวที่คัดกรองตามประเภทแล้วผ่านเว็บเซอร์วิส งานวิจัย [7] ของ Goossen และคณะได้กล่าวถึงการแนะนำข่าวสำหรับผู้ใช้งานโดยใช้หลักการ CF-IDF วิธีนี้เป็นการปรับจากวิธี TF-IDF (Term Frequency-Inverse Document Frequency) ซึ่งเป็นวิธีที่ใช้ในการกำหนดความสำคัญให้กับคำทุกคำในเอกสาร โดยเปลี่ยนมาพิจารณาความสำคัญของคำในเอกสารข่าวเฉพาะที่ตรงกับคอนเซปต์ในออนโทโลยีข่าวนั้น แล้วจึงเปรียบเทียบข่าวที่ตรงกับคำศัพท์ที่ระบุในโปรไฟล์ของผู้ใช้ การทดลองในงานวิจัยชิ้นนี้พบว่า CF-IDF มีประสิทธิภาพดีกว่า TF-IDF งานวิจัย [8] ของ Agarwal และคณะได้พัฒนาออนโทโลยีด้านการศึกษาจากคำในหมวด Education ในมาตรฐานไอพีทีซี เพื่อประโยชน์ในการนำเสนอข่าวด้านการศึกษา นอกจากนั้น Agarwal และคณะยังเสนองานวิจัย [4] ที่ทำการจำแนกประเภทข่าวโดยใช้วิธี Weighted CF-IDF ซึ่งปรับจาก CF-IDF ของงานวิจัย [7] มาใช้ในการคำนวณความสำคัญของคอนเซปต์ในออนโทโลยีซึ่งปรากฏในแต่ละข่าว โดยให้น้ำหนักความสำคัญเพิ่มแก่คอนเซปต์ที่ปรากฏทั้งที่หัวข้อข่าวและเนื้อข่าวมากกว่าคอนเซปต์ที่

ปรากฏที่เนื้อหาอย่างเดียวนั้นจึงทำการจำแนกข่าว จาก การทดลองวิธี Weighted CF-IDF มีประสิทธิภาพดีกว่า CF-IDF การนำเสนอข่าวในงานวิจัยที่กล่าวมาข้างต้นยังนำเสนอใน รูปของรายการข่าวเรียงตามลำดับ แต่ NewsMap [9] ซึ่งเป็น เว็บไซต์รวบรวมข่าวจากคุณกิลนิวส์ มีการนำเสนอข่าวโดยการ แสดงภาพด้วยเทคนิคทรีแมป (Treemap) สำหรับแสดงข้อมูลที่เป็น ลำดับชั้น (Hierarchy) ในรูปของกล่องสี่เหลี่ยมซ้อนกัน โดยสีแสดงประเภทข่าว ความเข้มของสีแสดงอายุข่าว และ ขนาดของกล่องแสดงปริมาณข่าวอื่นที่เกี่ยวข้อง ดังในภาพที่ 2 แนวทางของ NewsMap มีความน่าสนใจ ทำให้ผู้ใช้เห็น ภาพรวมและเข้าใจลักษณะของข่าว แต่ยังไม่สะดวกนักหาก ต้องการมองภาพพัฒนาการของข่าวเรื่องหนึ่ง ๆ



ภาพที่ 2: ตัวอย่างข่าวใน NewsMap

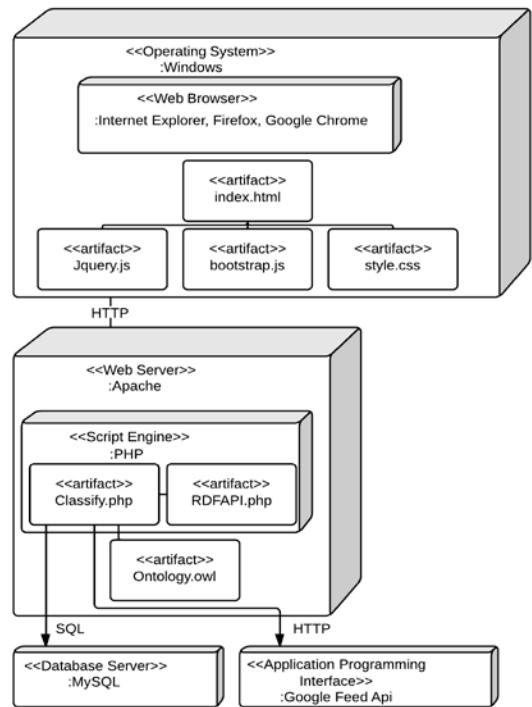
3. เครื่องมือแสดงผลภาพใหม่ไ้สำหรับข่าว

หัวข้อนี้จะอธิบายภาพรวมของเครื่องมือและรายละเอียด ขององค์ประกอบสำคัญได้แก่ การสร้างออนโทโลยีประเภท ข่าว การจำแนกประเภทข่าว และการแสดงผลแบบใหม่ไ้

3.1 ภาพรวมของเครื่องมือ

เครื่องมือแสดงผลภาพใหม่ไ้สำหรับข่าวนั้นพัฒนาขึ้นใน รูปแบบของเว็บแอปพลิเคชันตามสถาปัตยกรรมของเครื่องมือ ในภาพที่ 3 โดยใช้ Google Feed API [10] ในการดึงข้อมูลข่าว จากแหล่งข่าวทั้งต่างประเทศและในประเทศซึ่งมีการจัดเตรียม บริการการแบ่งปันข้อมูลข่าวระหว่างเว็บในรูปแบบอาร์เอสเอส เป็นภาษาอังกฤษ จำนวน 5 แหล่งข่าว เป็นกรณีศึกษา ได้แก่ 1) CNN 2) BBC 3) Daily Mail 4) Bangkok Post 5) The Nation แล้วเก็บข้อมูลข่าวที่ดึงมาไว้ในฐานข้อมูล และใช้ RDF API for PHP [11] ในการสกัดคอนเซ็ปต์จากออนโทโลยีประเภทข่าว (ดู หัวข้อที่ 3.2) เพื่อมาใช้ในการจำแนกข่าว ส่วนการแสดงผลใช้

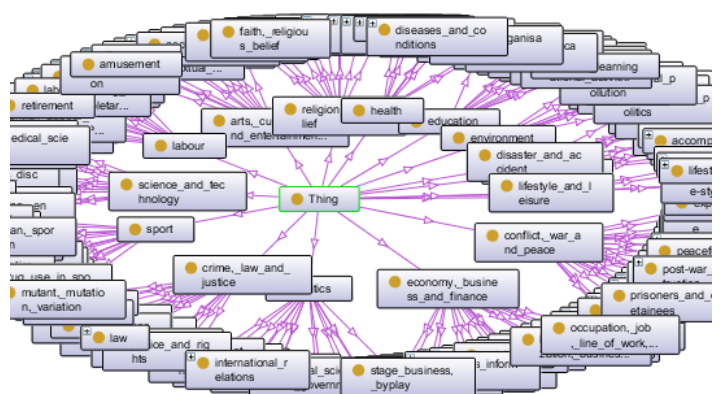
Jquery [12] และ bootstrap [13] ในการจัดการและควบคุมหน้า เว็บ รวมทั้งใช้ CSS (Cascading Style Sheets) ในการควบคุม การแสดงผลของหน้าเว็บให้อยู่ในรูปแบบที่เหมาะสม



ภาพที่ 3: สถาปัตยกรรมของเครื่องมือ

3.2 การสร้างออนโทโลยีประเภทข่าว

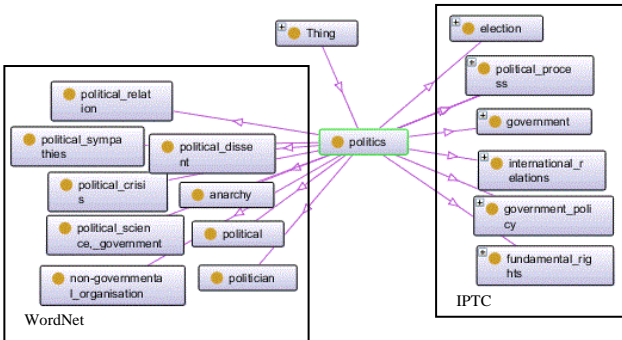
ผู้วิจัยสร้างออนโทโลยีสำหรับโดเมนประเภทข่าวโดยใช้ โปรแกรม Protégé [14] ดังภาพที่ 4 โดยมีจำนวนคำศัพท์หรือ คอนเซ็ปต์รวม 1,337 คำ



ภาพที่ 4: ออนโทโลยีประเภทข่าว

คำศัพท์ในออนโทโลยีประกอบด้วยสองส่วน ส่วนแรกเป็น คอนเซ็ปต์ประเภทข่าวจำนวน 17 ประเภทที่ระบุโดยมาตรฐาน ไอพีทีซี ได้แก่ 1) Arts, Culture and Entertainment 2) Crime,

Law and Justice 3) Disaster and Accident 4) Economy, Business and Finance 5) Education 6) Environment 7) Health 8) Human Interest 9) Labor 10) Lifestyle and Leisure 11) Politics 12) Religion and Belief 13) Science and Technology 14) Society 15) Sport 16) Conflicts, War and Peace 17) Weather จำนวนรวม 1,150 คำ ส่วนที่สองเป็นคอนเซปต์เสริมจากเวิร์ดเน็ตจำนวน 187 คำ ดังตัวอย่างในภาพที่ 5



ภาพที่ 5: ตัวอย่างคอนเซปต์ในประเภท politics

3.3 การจำแนกประเภทข่าว

ในการจำแนกประเภทข่าว ต้องทำการประมวลผลข้อความข่าวจากอาร์เอสเอสฟีดก่อน ขั้นตอนประกอบด้วย 1) การตัดคำจากข้อความ (Tokenization) 2) การลบคำที่ไม่มีนัยสำคัญออก (Stop-Word Removal) เช่น เครื่องหมายคำพูด คำเชื่อม เป็นต้น 3) สเต็มมิง (Stemming) เป็นการลดรูปของคำให้เป็นรากศัพท์ของคำนั้น ๆ เพื่อที่จะสามารถนำไปเปรียบเทียบกับคอนเซปต์ที่อยู่ในออนโทโลยีประเภทข่าวได้

การจำแนกประเภทข่าวจะใช้วิธี Weighted CF-IDF [4] กำหนดให้ออนโทโลยีประเภทข่าว C คือเซตของ t คอนเซปต์:

$$C = \{c_1, c_2, \dots, c_t\} \text{ where } t = 1,337 \quad (1)$$

แต่ละประเภทข่าว d_j คือเซตของ q คอนเซปต์ และข้อมูลข่าวฟีด r หนึ่งชิ้นคือเซตของ s คอนเซปต์:

$$d_j = \{c_1^{d_j}, c_2^{d_j}, \dots, c_q^{d_j}\}, \text{ where } c^{d_j} \in C \text{ and } j = 1, \dots, 17 \quad (2)$$

$$r = \{c_1^r, c_2^r, \dots, c_s^r\}, \text{ where } c^r \in C \quad (3)$$

เมื่อพิจารณาแต่ละประเภทข่าวตามออนโทโลยีและข้อมูลข่าวฟีดเป็นเอกสารที่ต้องเปรียบเทียบความคล้ายคลึงกันเพื่อการจำแนก เราจะใช้ทฤษฎี Vector Space Model ในการสร้างเวกเตอร์ที่เป็นตัวแทนเอกสาร

กำหนดให้เวกเตอร์ V_{d_j} เป็นตัวแทนเอกสารประเภทข่าว d_j :

$$V_{d_j} = (w_{1,j}, w_{2,j}, \dots, w_{t,j}) \quad (4)$$

โดยที่ w_{ij} คือค่าน้ำหนักความสำคัญของคอนเซปต์ c_i ใน d_j

โดยที่ $c_i \in C$ และ $w_{ij} \geq 0$ ค่า w_{ij} คำนวณโดย

$$w_{i,j} = \begin{cases} cf_{i,j} \times idf_i = \frac{f_{i,j}}{\sum_k f_{k,j}} \times \log \frac{N}{n_i} & \text{if } f_{i,j} > 0 \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

โดยที่ $cf_{i,j}$ = ค่าความถี่ของคอนเซปต์ (Concept Frequency) ของ c_i ในเอกสาร d_j

idf_i = ค่าความถี่ในเอกสารแบบผกผัน (Inverse Document Frequency) ของ c_i

f_{ij} = จำนวนครั้งที่คอนเซปต์ c_i ปรากฏในเอกสาร d_j

f_{kj} = จำนวนครั้งที่คอนเซปต์ c_k ใด ๆ ปรากฏใน

เอกสาร d_j

N = จำนวนเอกสารทั้งหมด และ

n_i = จำนวนเอกสารที่คอนเซปต์ c_i ปรากฏ

จากนั้นกำหนดให้เวกเตอร์ V_r เป็นตัวแทนเอกสารข่าวฟีด r :

$$V_r = (\alpha w_{1,r}, \alpha w_{2,r}, \dots, \alpha w_{t,r}) \quad (6)$$

โดยที่ w_{ir} คือค่าน้ำหนักความสำคัญของคอนเซปต์ c_i ใน r ซึ่งคำนวณแบบเดียวกับ w_{ij} โดยใช้ (5) และ

α คือค่าสัมประสิทธิ์สำหรับให้น้ำหนักความสำคัญเพิ่มแก่คอนเซปต์ c_i ในกรณีที่ c_i ปรากฏทั้งในหัวข้อข่าวและเนื้อข่าว โดย $\alpha > 1$; ถ้าไม่เช่นนั้นให้ c_i มี $\alpha = 1$

เมื่อได้เวกเตอร์ทั้งหมดแล้ว การจำแนกข่าวฟีดหนึ่งชิ้นว่าเป็นข่าวประเภทใด ทำโดยคำนวณค่าความคล้ายคลึงระหว่างเอกสารประเภทข่าวแต่ละประเภทกับเอกสารข่าวฟีด โดยใช้การวัดค่าความคล้ายคลึงเชิงมุม (Cosine Similarity):

$$sim(d_j, r) = \frac{V_{d_j} \cdot V_r}{|V_{d_j}| \times |V_r|} = \frac{\sum_{i=1}^t w_{i,j} \times w_{i,r}}{\sqrt{\sum_{i=1}^t w_{i,j}^2} \times \sqrt{\sum_{i=1}^t w_{i,r}^2}} \quad (7)$$

ค่า $sim(d_j, r)$ อยู่ระหว่าง 0 ถึง 1 ข้อมูลข่าวจะถูกจำแนกอยู่ในประเภทข่าว d_j ที่มีค่าความคล้ายคลึงนี้มากที่สุด

3.4 การแสดงภาพแบบใหม่ไลน์

ในงานวิจัยนี้นำเทคนิคการแสดงผลภาพ (Visualization) ในรูปแบบใหม่ไลน์มาใช้ในการนำเสนอข้อมูลข่าวแบบจำแนกตามประเภท เพื่อแสดงให้เห็นลำดับของข่าวที่เกิดขึ้น ภาพที่ 6

แสดงตัวอย่างหน้าจอใหม่ไลน์ของข่าว ซึ่งข่าวแต่ละประเภทจะแสดงด้วยสีที่แตกต่างกัน ผู้ใช้สามารถคลิกเปิดดูข้อมูลสี่ประจำประเภทข่าวได้คั้งด้านซ้ายของหน้าจอ ส่วนด้านขวาแสดงกล่องสรุปจำนวนข่าวในแต่ละประเภทในวันที่ปรากฏ โดยขนาดของกล่องสะท้อนถึงปริมาณข่าว

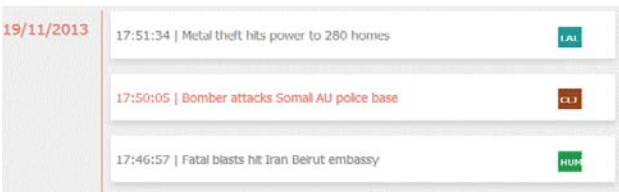


ภาพที่ 6: ตัวอย่างหน้าจอใหม่ไลน์

ในการแสดงภาพแบบใหม่ไลน์สำหรับข่าว ผู้ใช้สามารถระบุเงื่อนไข 1) เลือกช่วงวันหรือเวลา 2) เลือกมุมมอง 3) กรองออก 4) ค้นหาคำสำคัญ ทั้งนี้เมื่อมีการแสดงผลข่าวตามเงื่อนไขเหล่านี้แล้ว ผู้ใช้สามารถคลิกดูรายละเอียดของข่าวได้

การเลือกช่วงวันหรือเวลาเป็นการเลือกเพื่อเข้าถึงข้อมูลข่าวเฉพาะในวันหรือเวลาที่เลือก เมื่อเลือกวัน เครื่องมือจะแสดงผลเฉพาะข้อมูลข่าวที่เกิดขึ้นเฉพาะวันนั้น และแบ่งข้อมูลข่าวออกเป็นช่วงเวลา แต่หากเลือกเวลา เครื่องมือจะแสดงผลข้อมูลข่าวเฉพาะช่วงเวลาที่ทำกรเลือก

การเลือกมุมมองเป็นการเลือกว่าจะแสดงผลข่าวโดยการจัดกลุ่ม ดังเช่นด้านขวาของภาพที่ 6 หรือจะแสดงผลข่าวเป็นลิสต์รายการแทน ดังภาพที่ 7

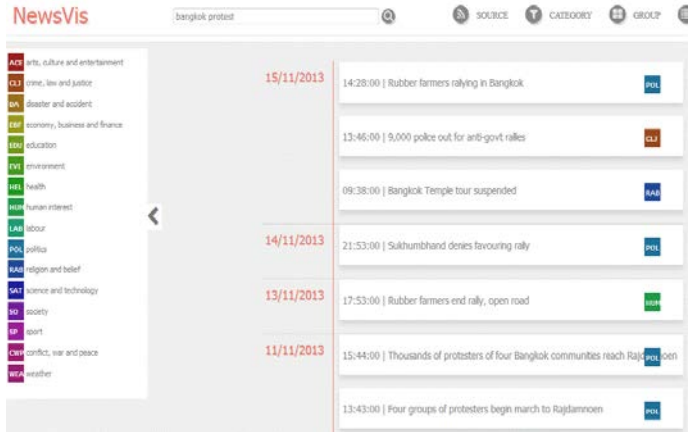


ภาพที่ 7: ตัวอย่างมุมมองแบบลิสต์รายการ

การกรองออกเป็นการเลือกกรองประเภทข่าวหรือแหล่งข่าวที่ไม่ต้องการออกจากการแสดงผล

การค้นหาคำสำคัญเป็นการค้นหาเฉพาะข้อมูลข่าวที่มีคำสำคัญนั้น ๆ ปรากฏอยู่ เพื่อแสดงลำดับและช่วงเวลาที่เกิดขึ้น

ของข้อมูลข่าวที่มีคำสำคัญนั้นปรากฏอยู่ ดังภาพที่ 8 แสดงถึงลำดับของข้อมูลข่าวที่มีคำว่า bangkok protest ปรากฏอยู่ ทำให้สามารถเห็นภาพพัฒนาการของข่าวที่เกี่ยวข้องกับคำสำคัญนี้ได้



4. การประเมินผลและอภิปรายผล

ผู้วิจัยรวบรวมข้อมูลอาร์เอสเอสทีดจาก 5 แหล่งข่าวตามทีระบุในหัวข้อที่ 3.1 ในระหว่างวันที่ 10-19 พฤศจิกายน 2556 จำนวนรวม 1,250 ข่าว จากนั้นทำการประเมินผลการจำแนกโดยใช้ข้อมูลข่าวจำนวน 100 ข่าว (20 ข่าวล่าสุดจาก 5 แหล่งข่าว) โดยเปรียบเทียบผลการจำแนกโดยเครื่องมือกับการจำแนกด้วยมือ และใช้ค่าความแม่นยำ (Precision) ค่าเรียกคืน (Recall) กับค่าเอฟ-เมเชอร์ (F-measure) ซึ่งเป็นค่าประสิทธิภาพโดยรวมของค่าความแม่นยำและค่าเรียกคืน เป็นตัววัด:

$$precision = \frac{\text{number of news feeds correctly classified by tool}}{\text{number of news feeds classified by tool}} \quad (8)$$

$$recall = \frac{\text{number of news feeds correctly classified by tool}}{\text{number of news feeds classified manually}} \quad (9)$$

$$F\text{-measure} = 2 * \frac{precision * recall}{precision + recall} \quad (10)$$

ผลการประเมินในตารางที่ 1 เมื่อทดลองกำหนด α ใน (6) เป็นค่าต่าง ๆ ได้ผลว่า มีค่าความแม่นยำและค่าเรียกคืนค่อนข้างน่าพอใจ โดยเมื่อ $\alpha = 1.6$ จะได้ผลดีที่สุดโดยมีค่าความแม่นยำเป็น 72.52% ค่าเรียกคืนเป็น 66% และค่าเอฟ-เมเชอร์เป็น 69.11% ค่า α นี้จึงเป็นค่าที่ใช้ในเครื่องมือ

จากการที่ข้อมูลข่าวอาร์เอสเอสทีดจะเป็นสรุปข่าว การประมวลผลทำกับเพียงหัวข้อและเนื้อข่าวที่สั้น การเพิ่มค่าน้ำหนักให้กับคอนเซปต์ที่พบในข่าวด้วยค่าสัมประสิทธิ์ α ที่

แตกต่างกัน จึงไม่มีผลแตกต่างกันมากนักเนื่องจากคอนเซปต์ที่เกี่ยวข้องกับข่าวและจะถูกกระทบจากการปรับค่าน้ำหนักด้วยค่าสัมประสิทธิ์นั้นมีน้อย และจากการที่ปริมาณเนื้อหาข่าวที่ใช้วิเคราะห์มีน้อย ผู้วิจัยจึงเห็นว่าผลการประเมินที่ให้ค่าเอฟ-เมเชอร์มากกว่า 65% จึงสะท้อนประสิทธิภาพโดยรวมที่ค่อนข้างสูงและน่าพอใจ อย่างไรก็ตาม การจำแนกประเภทข่าวยังมีข้อจำกัดอื่นจากการจำแนกด้วยคำจำนวนน้อยในข่าว เช่น ในภาพที่ 8 ข่าวเกี่ยวกับ Rubber farmers สองชิ้นถูกจำแนกเป็นคนละประเภท เนื่องจากชิ้นที่ถูกจัดเป็น human interest กล่าวถึงสั้น ๆ เรื่องการยกเลิกปิดถนน โดยไม่มีข้อความมากพอที่จะบอกบริบทของข่าวได้ว่าเกี่ยวข้องกับการเมือง ในทำนองเดียวกันจากภาพที่ 8 ข่าว Bangkok temple tour ถูกจำแนกเป็นประเภท religion and belief เนื่องด้วยการให้ความสำคัญกับคำว่า temple ที่ปรากฏทั้งในหัวข้อข่าวและเนื้อหาข่าวโดยไม่มีข้อความมากพอที่แสดงน้ำหนักในบริบททางด้านการเมือง

ตารางที่ 1: ผลการประเมินตามค่าสัมประสิทธิ์ α

ค่าสัมประสิทธิ์ α	ค่าความแม่นยำ	ค่าเรียกคืน	ค่าเอฟ-เมเชอร์
1.0	71.42%	65%	68.06%
1.2	70.32%	64%	67.01%
1.4	70.32%	64%	67.01%
1.6	72.52%	66%	69.11%
1.8	71.42%	65%	68.06%
2.0	71.42%	65%	68.06%

5. บทสรุป

เครื่องมือแสดงภาพใหม่ไคลน์สำหรับข่าวนี้เป็นทางเลือกช่วยเสริมนอกเหนือจากการรับข่าวสารจากเว็บไซต์ข่าวโดยตรง โดยให้ข้อมูลแบบสรุปข่าวที่ช่วยจำแนกและแสดงผลข่าวจากหลายแหล่งข่าวให้กับผู้ใช้ในหลายมุมมอง รวมทั้งมุมมองของพัฒนาการของข่าวที่สนใจ ในการปรับปรุงเครื่องมือนี้เพิ่มเติมสามารถเพิ่มการแสดงผลแบบอื่นร่วมด้วย เช่น การแสดงผลสรุปข่าวอิงสถานที่ที่เกิดข่าว การใช้วิธีการจำแนกที่มี

ประสิทธิภาพมากขึ้น เช่น เอสวีเอ็ม (SVM) รวมไปถึงการขยายออนโทโลยีให้ครอบคลุมโดเมนข่าวมากขึ้น

เอกสารอ้างอิง

- [1] W3C, *RSS Tutorial*, <http://www.w3schools.com/rss/>, Accessed date: August 15, 2013.
- [2] International Press Telecommunications Council, *News Exchange Format Standard*, <http://www.iptc.org/>, Accessed date: August 10, 2013.
- [3] Princeton Univ, *WordNet*, <http://wordnet.princeton.edu/>, Accessed date: August 18, 2013.
- [4] S. Agarwal, A. Singhal, and P. Bedi, "Classification of RSS feed news items using ontology," *Proc. 12th Int. Conf. Intelligent Systems Design and Applications (ISDA 2012)*, 27-29 Nov 2012, pp. 491-496.
- [5] Wikipedia, *Visualization (computer graphics)*, [http://en.wikipedia.org/wiki/Visualization\(computer_graphics\)](http://en.wikipedia.org/wiki/Visualization(computer_graphics)), Accessed date: September 17, 2013.
- [6] S. Saha, A. Sajjanhar, S. Gao, and R. Dew, "Delivering categorized news items using RSS feeds and Web services," *Proc. 10th IEEE Int. Conf. Computer and Information Technology (CIT 2010)*, 2010, pp. 699-712.
- [7] F. Goossen, W. IJntema, F. Frasincar, F. Hogenboom, and U. Kaymak, "News personalization using the CF-IDF semantic recommender," *Proc. Int. Conf. Web Intelligence, Mining and Semantics (WIMS '11)*, 2011.
- [8] S. Agarwal, A. Singhal, and P. Bedi, "IPTC based ontological representation of educational news RSS feeds," *Proc. 3rd Int. Conf. Trends in Information, Telecommunication and Computing, Lecture Notes in Electrical Engineering, Volume 150*, 2013, pp. 353-359.
- [9] Marcos Weskamp, *NewsMap*, <http://newsmap.jp/>, Accessed date: November 11, 2013.
- [10] Google, *Google Feed API*, <https://developers.google.com/feed/>, Accessed date: September 17, 2013.
- [11] D. Westphal and C. Bizer, *RDF API for PHP*, <http://wifo5-03.informatik.uni-mannheim.de/bizer/rdxfapi/>, Accessed date: September 19, 2013.
- [12] JQuery, *Jquery*, <http://www.jquery.org/>, Accessed date: September 20, 2013.
- [13] Bootstrap, *Bootstrap*, <http://getbootstrap.com/>, Accessed date: September 20, 2013.

- [14] Stanford Center for Biomedical Informatics Research,
Protégé, <http://protege.stanford.edu/>, Accessed date:
August 20, 2013.

การสร้างกรณีทดสอบโดยใช้ขั้นตอนวิธีเชิงพันธุกรรม

Test Case Generation Using Genetic Algorithms

นันทนี ช่วยชู(Nuntanee Chuaychoo)¹ และ สุภาภรณ์ กานต์สมเกียรติ (Supaporn Kansomkeat)²

¹สาขาวิชาการจัดการเทคโนโลยีสารสนเทศ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยสงขลานครินทร์

²ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยสงขลานครินทร์

¹nun-ta-nee@hotmail.com, ²supaporn.k@psu.ac.th

บทคัดย่อ

การทดสอบซอฟต์แวร์ เป็นขั้นตอนหนึ่งที่สำคัญในกระบวนการผลิตซอฟต์แวร์ การทดสอบช่วยให้ซอฟต์แวร์ที่พัฒนามีความถูกต้อง และมีความน่าเชื่อถือเพิ่มมากขึ้น การสร้างกรณีทดสอบเป็นอีกองค์ประกอบหนึ่งที่มีความสำคัญในกระบวนการทดสอบซอฟต์แวร์ คุณภาพของการทดสอบจะขึ้นอยู่กับประสิทธิภาพของกรณีทดสอบ บทความนี้นำเสนอวิธีการสร้างกรณีทดสอบโดยใช้ขั้นตอนวิธีเชิงพันธุกรรม และได้นำขั้นตอนการสร้างกรณีทดสอบที่ได้นำเสนอไปประยุกต์ใช้กับกรณีศึกษา ผลลัพธ์ที่ได้พบว่า กรณีทดสอบที่สร้างขึ้นตามวิธีการที่นำเสนอมีความเหมาะสมสำหรับการทดสอบซอฟต์แวร์

คำสำคัญ: การทดสอบซอฟต์แวร์, ขั้นตอนวิธีเชิงพันธุกรรม, การสร้างกรณีทดสอบ

Abstract

Software Testing is an important step in the software process. It makes the developed software more accurate and reliable. Generating test cases is a key element in the process of software testing. The quality of testing depends on the effectiveness of test. This paper presents test case generation method using genetic algorithms. The proposed method was applied on a case study. The result shows that the generated test cases are appropriate for testing software.

Keywords : Software Testing, Genetic Algorithms, Test

Case Generation

1. บทนำ

ในปัจจุบันความต้องการของการใช้คอมพิวเตอร์และเทคโนโลยีสารสนเทศยังคงมีเพิ่มขึ้นเรื่อยๆ ดังนั้นจึงมีความต้องการซอฟต์แวร์เพื่อใช้งานเฉพาะต่าง ๆ อีกมากมาย การทดสอบซอฟต์แวร์ (Software Testing) [1] เป็นขั้นตอนที่สำคัญในกระบวนการพัฒนาซอฟต์แวร์ที่มีบทบาทและมีความสำคัญมาก การทดสอบเป็นกิจกรรมที่จัดทำขึ้นเพื่อปรับปรุงคุณภาพของซอฟต์แวร์ ค้นหาข้อบกพร่องและลดข้อผิดพลาดจากการทำงานของซอฟต์แวร์ให้เหลือน้อยที่สุด

การกำหนดกรณีทดสอบเป็นสิ่งหนึ่งที่สำคัญในกระบวนการทดสอบซอฟต์แวร์ ขั้นตอนการสร้างกรณีทดสอบที่ดี จะมีส่วนทำให้การทดสอบมีประสิทธิภาพ กรณีทดสอบที่มีประสิทธิภาพจะต้องครอบคลุมคุณลักษณะต่างๆ ของระบบและมีจำนวนกรณีทดสอบที่ไม่มากเกินไป

ขั้นตอนวิธีเชิงพันธุกรรม (Genetic Algorithms) เป็นวิธีการค้นหาคำตอบที่เหมาะสมโดยใช้หลักการคัดเลือกแบบธรรมชาติจากการจำลองแนวคิดวิวัฒนาการของสิ่งมีชีวิต การนำขั้นตอนวิธีทางพันธุกรรมมาประยุกต์ใช้ในการสร้างกรณีทดสอบสำหรับกระบวนการทดสอบซอฟต์แวร์ จึงเป็นเรื่องที่น่าสนใจ เนื่องจากหลักการทำงานของขั้นตอนวิธีทางพันธุกรรม มีความสอดคล้องกับขั้นตอนการค้นหาค่ากรณีทดสอบและมีการปรับเปลี่ยนกรณีทดสอบ จนกระทั่งได้กรณีทดสอบที่เหมาะสมที่สุด สามารถช่วยให้ผู้พัฒนาระบบหรือผู้ทดสอบระบบสามารถทำงานได้ง่ายขึ้น ช่วยลดเวลาและต้นทุนในขั้นตอนการทดสอบซอฟต์แวร์

บทความนี้แนะนำการสร้างกรณีทดสอบที่สามารถครอบคลุมการทำงานและเงื่อนไขการทำงานของโปรแกรม โดยประยุกต์ใช้ขั้นตอนวิธีเชิงพันธุกรรม

บทความนี้ประกอบด้วยหัวข้อต่างๆ ดังต่อไปนี้คือ หัวข้อที่ 2 เสนอทฤษฎีที่เกี่ยวข้อง หัวข้อที่ 3 เสนอวิธีการสร้างกรณีทดสอบโดยใช้ขั้นตอนวิธีเชิงพันธุกรรม หัวข้อที่ 4 เสนอกรณีศึกษา และหัวข้อที่ 5 สรุปผลและงานในอนาคต

2. ทฤษฎีที่เกี่ยวข้อง

ในส่วนนี้นำเสนอเกี่ยวกับทฤษฎีที่เกี่ยวข้อง ประกอบไปด้วยการทดสอบซอฟต์แวร์และขั้นตอนวิธีเชิงพันธุกรรม

2.1 การทดสอบซอฟต์แวร์ (Software Testing)

การทดสอบเป็นปัจจัยสำคัญปัจจัยหนึ่งของการประกันคุณภาพซอฟต์แวร์ ดังนั้นในกระบวนการพัฒนาซอฟต์แวร์จึงต้องมีขั้นตอนการทดสอบซอฟต์แวร์ สำหรับเทคนิคในการทดสอบซอฟต์แวร์ (Software Testing Technique) ที่นิยมใช้กันแพร่หลายได้แก่

1. Black-Box Testing คือการทดสอบการทำงานของซอฟต์แวร์ที่ไม่สนใจกลไกภายในของระบบ การทำการทดสอบเน้นเพื่อตรวจสอบผลการการทำงานของระบบในแต่ละหน้าที่ตามข้อกำหนดความต้องการ (Requirement Specification) ว่าถูกต้องตามความต้องการหรือไม่

2. White-Box Testing คือการทดสอบการทำงานของซอฟต์แวร์ ซึ่งพิจารณากลไกภายในของระบบโดยจะมุ่งเน้นพิจารณาโครงสร้างภายใน

2.2 ขั้นตอนวิธีเชิงพันธุกรรม (Genetic Algorithms)

ขั้นตอนวิธีเชิงพันธุกรรม เป็นกระบวนการหนึ่งที่น่าสนใจ ทฤษฎีการวิวัฒนาการของสิ่งมีชีวิต นำเสนอโดย John Holland [2] สามารถนำมาใช้เพื่อค้นหาคำตอบที่ดีที่สุดและเหมาะสมมากขึ้นโดยผ่านการเรียนรู้ด้วยตัวเอง เปรียบเสมือนการวิวัฒนาการของสิ่งมีชีวิตที่สามารถวิวัฒนาการเพื่อการดำรงชีวิตอยู่ กระบวนการทำงานของขั้นตอนวิธีเชิงพันธุกรรมจะเริ่มจากการสุ่มโครโมโซมขึ้นมาจำนวนหนึ่งเพื่อสร้างเป็นกลุ่มประชากรต้นกำเนิด (Initial Population) และทำการประเมินค่าความเหมาะสม (Fitness Evaluation) จากนั้นจะคัดเลือกประชากร (Selection) มาจำนวนหนึ่งเพื่อใช้เป็นต้นกำเนิดทางสายพันธุ์หรือเป็นกลุ่มประชากรรุ่นพ่อแม่ จากนั้นจะนำประชากรเหล่านี้เข้าสู่กระบวนการปฏิบัติการทางสายพันธุ์ (Genetic Operation) ได้แก่ การไขว้

เปลี่ยน (Crossover) และการกลายพันธุ์ (Mutation) ซึ่งจะทำให้เกิดกลุ่มประชากรที่เป็นรุ่นลูก และเมื่อนำประชากรกลุ่มนี้ไปประเมินค่าความเหมาะสมแล้วพบคำตอบที่เหมาะสมกับปัญหา กระบวนการทำงานก็จะจบลง แต่หากคำตอบที่ได้ยังไม่เหมาะสม ก็จะดำเนินต่อไปโดยการนำประชากรรุ่นลูกที่ได้ไปแทนที่ (Replacement) ประชากรรุ่นพ่อแม่ เมื่อทำการแทนที่เสร็จแล้วก็จะนำประชากรกลุ่มใหม่ที่ได้กลับเข้าสู่กระบวนการคัดเลือกใหม่อีกครั้ง ซึ่งกระบวนการเหล่านี้จะดำเนินการซ้ำไปเรื่อยๆจนกระทั่งได้ประชากรกลุ่มที่เหมาะสม

3. การสร้างกรณีทดสอบโดยใช้ขั้นตอนวิธีเชิงพันธุกรรม

ขั้นตอนการสร้างกรณีทดสอบโดยใช้ขั้นตอนวิธีเชิงพันธุกรรม แสดงดังภาพที่ 1 ประกอบด้วย 2 ขั้นตอนหลัก คือ

- 1) การสร้างกราฟควบคุมการไหลและระบุเส้นทางเป้าหมาย (Generating Control Flow Graph and Specifying Target Paths)
- 2) การสร้างกรณีทดสอบ (Generating Test Cases)

3.1 การสร้างกราฟควบคุมการไหลและระบุเส้นทางเป้าหมาย

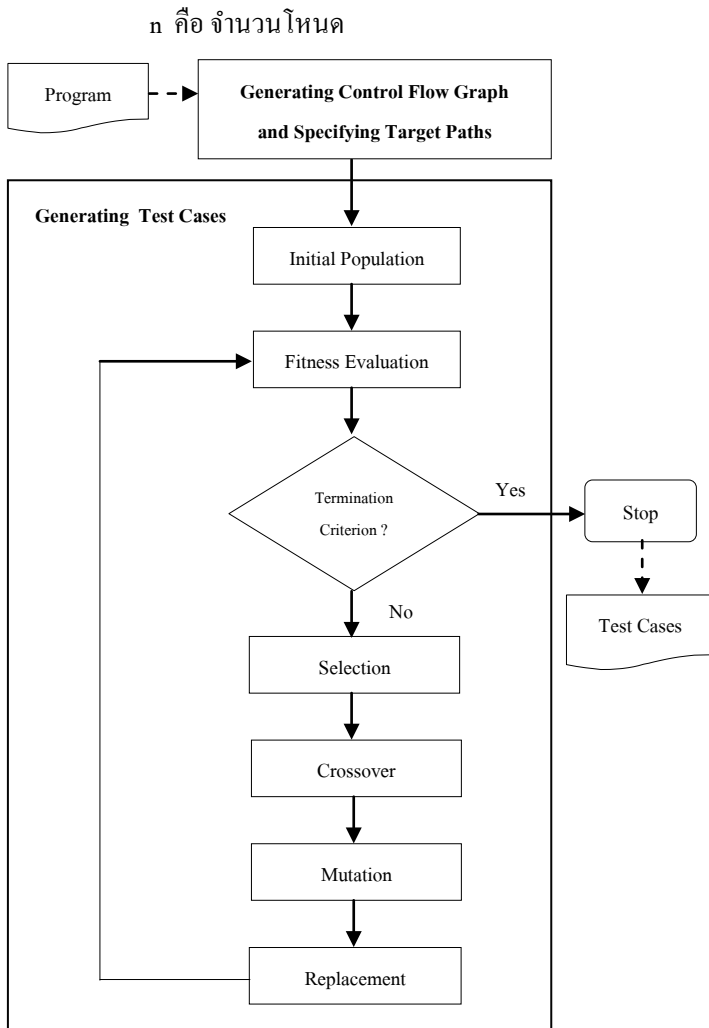
การสร้างกราฟควบคุมการไหล พิจารณาจากรหัสต้นทาง (Source Code) ของโปรแกรมที่ต้องการทดสอบ กราฟควบคุมการไหลสร้างขึ้นตามหลักการที่ได้นำเสนอโดย Amman และ Offutt [3] ภาพที่ 2 แสดงตัวอย่างรหัสต้นทางของโปรแกรม Triangle ที่ต้องการทดสอบ ภาพที่ 3 แสดงกราฟควบคุมการไหลของโปรแกรม Triangle

กราฟควบคุมการไหลจากขั้นตอนข้างต้น จะถูกนำมาใช้เพื่อระบุเส้นทางเป้าหมาย โดยจะคำนวณจำนวนเส้นทางจากการคำนวณค่าความซับซ้อนไซโคลมาติก (Cyclomatic Complexity) [4] ซึ่งใช้ในการวัดความซับซ้อนของโปรแกรมเป็นตัวบ่งชี้ว่าโปรแกรมนั้นมีความยากต่อการบำรุงรักษาและการทดสอบมากน้อยเพียงใด โดยสามารถคำนวณได้ดังนี้

$$V(G) = e - n + 2 \quad (1)$$

โดยที่ $V(G)$ คือ ค่าความซับซ้อนไซโคลมาติก

e คือ จำนวนเส้นเชื่อมต่อระหว่างโหนด



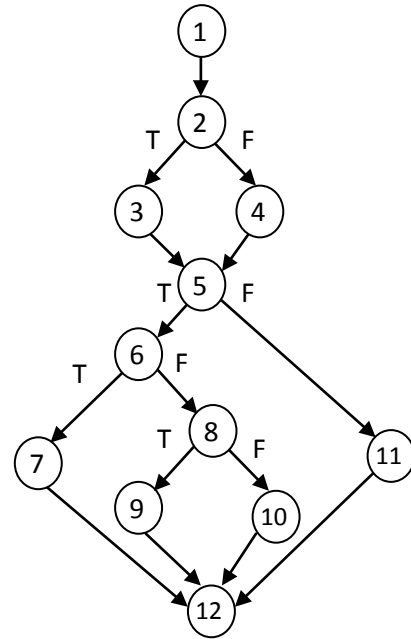
ภาพที่ 1 : แผนภาพขั้นตอนการสร้างกรณีทดสอบโดยใช้ขั้นตอนวิธีเชิงพันธุกรรม

```

1  Input (a,b,c) as Integer
2  If (a<b + c) And (b<a + c) And (c<b + a) Then
3    Is_Triangle = True
4  Else
5    Is_Triangle = False
6  End If
7  If Is_Triangle Then
8    If (a = b) And (b = c) Then
9      output = "Equilateral"
10   ElseIf (a<>b) And (b<>c) And (a<>c) Then
11     output = "Scalene"
12   Else
13     output = "Isosceles"
14   End If
15 Else
16   output = "Not a Triangle"
17 End If

```

ภาพที่ 2 : แสดงตัวอย่างรหัสต้นทางของโปรแกรม Triangle



ภาพที่ 3 : แสดงกราฟควบคุมการไหลของโปรแกรม Triangle

3.2 การสร้างกรณีทดสอบ

การสร้างกรณีทดสอบถูกดำเนินการโดยใช้กระบวนการขั้นตอนวิธีเชิงพันธุกรรม ประกอบด้วยขั้นตอนดังนี้
ขั้นตอนที่ 1 การสร้างกลุ่มประชากรต้นกำเนิด (Initial Population)

กลุ่มประชากรต้นกำเนิดจะถูกสร้างขึ้นโดยใช้วิธีการแบ่งส่วนที่สมดุล (Equivalence Partitioning) ข้อมูลนำเข้าของโปรแกรมที่ต้องการทดสอบแต่ละตัวจะถูกแบ่งส่วนออกเป็นกลุ่มๆ ประชากรต้นกำเนิดหนึ่งประชากร ได้จากการนำส่วนใดส่วนหนึ่งของข้อมูลนำเข้าทุกตัวมารวมกัน ดังนั้นหนึ่งประชากรต้นกำเนิดหมายถึง หนึ่งชุดข้อมูลทดสอบ ซึ่งจำนวนโครโมโซมของหนึ่งประชากรจะเท่ากับจำนวนตัวแปรของหนึ่งชุดข้อมูลทดสอบ

ขั้นตอนที่ 2 การประเมินค่าความเหมาะสม (Fitness Evaluation)

การประเมินค่าความเหมาะสมจะพิจารณาจากการประเมินค่าความครอบคลุมของชุดข้อมูลที่ใช้ในการทดสอบสำหรับบทความนี้ การประเมินค่าความครอบคลุมใช้วิธีการคำนวณหาค่า distance ของคำสั่งที่เป็นเงื่อนไขในการทำงานของโปรแกรม ตามทฤษฎี Korel's distance function [5] ดังแสดงในตารางที่ 1

ตารางที่ 1 : Korel's distance function

No	Predicate	Distance if path taken
1	A=B	ABS(A-B)
2	A≠B	K
3	A<B	(A-B)+K
4	A≤B	(A-B)
5	A>B	(B-A)+K
6	A≥B	(B-A)
7	X OR Y	MIN(Distance(X),Distance(Y))
8	X AND Y	(Distance(X) + Distance(Y)

ค่าความเหมาะสม สามารถคำนวณจากค่าผลต่างระหว่าง distance ดังนี้

$$\text{Fitness function} = \{ \text{dist(TG)} \} - \{ \text{dist(TR)} \} \quad (2)$$

โดยที่ dist(TG) คือ distance ของเส้นทางเป้าหมาย (Target Path)

dist(TR) คือ distance ของเส้นทางการเดินทางที่เกิดจากชุดข้อมูลทดสอบ (Traversal Path)

ค่าความเหมาะสมที่ได้ สามารถประเมินค่าความครอบคลุมได้ ดังนี้

$$\text{Fitness function} = \begin{cases} 0 & ; \text{covered} \\ \text{otherwise} & ; \text{not covered} \end{cases}$$

กรณีที่ค่าความเหมาะสม มีค่าเท่ากับศูนย์ แสดงว่าชุดข้อมูลทดสอบนั้นสามารถใช้เป็นกรณีทดสอบสำหรับเส้นทางเป้าหมายที่กำลังพิจารณา ในกรณีที่ค่าความเหมาะสม มีค่าไม่เท่ากับศูนย์ แสดงว่าชุดข้อมูลทดสอบนั้นยังไม่ครอบคลุมเส้นทางเป้าหมาย

เส้นทางเป้าหมายต่างๆ จะถูกตรวจสอบ โดยจะทำการตรวจสอบว่าชุดข้อมูลต่างๆ ครอบคลุมทุกเส้นทางเป้าหมายหรือไม่ กรณีที่ครอบคลุมทุกเส้นทางจะสิ้นสุดการทดสอบ

ขั้นตอนที่ 3 การคัดเลือก (Selection)

การคัดเลือกจะกระทำการคัดเลือกชุดข้อมูลเพื่อนำมาเป็นประชากรรุ่นพ่อแม่ โดยเลือกชุดข้อมูลที่ให้ค่าความเหมาะสม

น้อย ซึ่งหมายถึงชุดข้อมูลที่ให้ค่าผลต่างระหว่าง distance น้อยที่สุด

ขั้นตอนที่ 4 การไขว้เปลี่ยน (Crossover)

การไขว้เปลี่ยนจะดำเนินการโดยเลือกชุดข้อมูลที่ให้ค่าผลต่างระหว่าง distance น้อยที่สุดมา 2 ชุดข้อมูล และทำการไขว้เปลี่ยน โดยใช้วิธีการไขว้เปลี่ยนแบบจุดเดียว (Single-Point Crossover) ซึ่งจะทำให้ได้ชุดข้อมูลใหม่ 2 ชุดข้อมูล ตัวอย่างเช่น ชุดข้อมูลที่ถูกเลือกมาคือ $\{x_1, x_2, x_3, x_4\}$ และ $\{y_1, y_2, y_3, y_4\}$ จะทำการไขว้เปลี่ยนชุดข้อมูลทั้งสอง ณ ตำแหน่งที่ 4 ดังนั้นจะได้ชุดข้อมูลใหม่ คือ $\{x_1, x_2, x_3, y_4\}$ และ $\{y_1, y_2, y_3, x_4\}$

ขั้นตอนที่ 5 การกลายพันธุ์ (Mutation)

การกลายพันธุ์จะกระทำโดยเลือกชุดข้อมูลที่ให้ค่าผลต่างระหว่าง distance น้อยที่สุด เพื่อทำการกลายพันธุ์แบบจุดเดียว (Single-Point Mutation) ซึ่งจะทำให้ได้ชุดข้อมูลใหม่ 1 ชุดข้อมูล ตัวอย่างเช่น ชุดข้อมูลที่ถูกเลือกมาคือ $\{x_1, x_2, x_3, x_4\}$ จะทำการกลายพันธุ์ ณ ตำแหน่งที่ 2 ด้วยค่า x_m ดังนั้นจะได้ชุดข้อมูลใหม่ คือ $\{x_1, x_m, x_3, x_4\}$

โดยที่ตำแหน่งหรือโครโมโซมที่จะกระทำการไขว้เปลี่ยนในขั้นตอนที่ 4 และการกลายพันธุ์ในขั้นตอนที่ 5 จะได้จากการคัดเลือกโดยวิธีการสุ่ม เนื่องจากจำนวนตัวแปรหรือโครโมโซมของชุดข้อมูลมีจำนวนไม่มาก ดังนั้นการดำเนินการไขว้เปลี่ยนและการกลายพันธุ์จึงมีการกระทำแบบจุดเดียว

ขั้นตอนที่ 6 การแทนที่ (Replacement)

ชุดข้อมูลชุดใหม่ที่ได้จากขั้นตอนที่ 4 และ 5 จะถูกนำไปแทนที่ข้อมูลที่เป็นประชากรรุ่นพ่อแม่ และนำข้อมูลชุดใหม่ไปประเมินค่าความเหมาะสมในขั้นตอนที่ 2 อีกครั้ง

การกระทำขั้นตอนที่ 1 – 6 จะถูกกระทำซ้ำๆ โดยจำนวนรุ่นของประชากรจะเท่ากับจำนวนรอบการทำงานซ้ำ โดยกำหนดให้มีการทำซ้ำสูงสุด 20 รอบ

4. กรณีศึกษา

บทความนี้เลือกใช้โปรแกรม Triangle เพื่อแสดงการประยุกต์ใช้ขั้นตอนการหาค่าที่ได้นำเสนอ

จากกราฟควบคุมการไหลของโปรแกรม Triangle ดังแสดงในภาพที่ 3 สามารถคำนวณหาความซับซ้อนไซโคลมาติกได้เท่ากับ 5 ดังนั้นเส้นทางในการทดสอบโปรแกรมมี 5 เส้นทางซึ่งสามารถระบุเส้นทางเป้าหมายทั้ง 5 เส้นทางได้ดังตารางที่ 2

ตารางที่ 2 : เส้นทางสำหรับทดสอบโปรแกรม Triangle

Path	The description of target path
P1	1-2T-3-5T-6T-7-12
P2	1-2T-3-5T-6F-8T-9-12
P3	1-2T-3-5T-6F-8F-10-12
P4	1-2F-4-5F-11-12
P5	1-2T-3-5F-11-12

กลุ่มประชากรต้นกำเนิด ซึ่งเป็นกลุ่มของชุดข้อมูลเริ่มต้นจะถูกสร้างขึ้นโดยใช้วิธีแยกส่วนที่สมดุล (Equivalence Partitioning) สำหรับโปรแกรม Triangle มีข้อมูลนำเข้าเป็นจำนวนเต็ม 3 จำนวน (a, b, c) โดยจำนวนเต็มสามารถแบ่งส่วน (Partition) ของข้อมูลออกเป็น 3 กลุ่ม คือ มากกว่า, น้อยกว่า และเท่ากับศูนย์

การกำหนดกลุ่มของชุดข้อมูลเริ่มต้น จะกำหนดตามหลักการ Each Choice (EC) ที่ถูกนำเสนอโดย Amman และ Offutt [3] โดยหลักการนี้กำหนดให้แต่ละส่วนที่ถูกแบ่ง จะต้องถูกใช้อย่างน้อย 1 ครั้งในการทดสอบ ดังนั้น ชุดข้อมูลที่ใช้สำหรับโปรแกรม Triangle จึงประกอบด้วย 3 ชุดข้อมูล คือ $\{a<0,b<0,c<0\}$, $\{a=0,b=0,c=0\}$ และ $\{a>0,b>0,c>0\}$ จากนั้นทำการสุ่มค่าของข้อมูลที่ใช้สำหรับทดสอบแสดงดังตารางที่ 3

ตารางที่ 3 : ชุดข้อมูลที่เบื้องต้น สำหรับการทดสอบโปรแกรม

Triangle

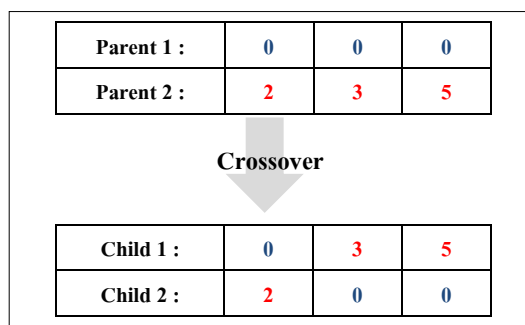
ชุดข้อมูล ตัวแปร	a	b	c
1	-3	-2	-5
2	0	0	0
3	2	3	5

การประเมินค่าความครอบคลุมของชุดข้อมูลที่ใช้ในการทดสอบถูกกระทำโดยวิธีการคำนวณค่า distance ของค่าสิ่งที่เป็นเงื่อนไขตามทฤษฎี Korel's distance function โดยเส้นทางเป้าหมายแต่ละเส้นทางจะถูกนำมาพิจารณาเพื่อหาชุดข้อมูลที่เหมาะสมสำหรับการทดสอบเส้นทางนั้นๆ ในกรณีที่ยังไม่มีชุดข้อมูลที่เหมาะสมสำหรับเส้นทางใด เส้นทางนั้นจะถูกดำเนินการพิจารณาต่อไป

ตัวอย่างเช่น จากตารางที่ 2 กำหนดให้ P4 เป็นเส้นทางเป้าหมายสำหรับการหาค่าที่ได้นั้น ดังนั้นค่าสิ่งที่เป็นเงื่อนไขในเส้นทาง P4 ประกอบด้วย $\{2F, 5F\}$ พิจารณาชุดข้อมูลทั้ง 3 ในตารางที่ 3 นำแต่ละชุดข้อมูลมาคำนวณหาผลต่างระหว่าง distance ของเส้นทางการเดินทางของข้อมูลชุดนั้น กับเส้นทางเป้าหมาย P4 จะพบว่า ผลต่างระหว่าง distance ของเส้นทางการเดินทางของข้อมูลชุดที่ 1 กับเส้นทางเป้าหมาย P4 มีค่าเท่ากับศูนย์ แสดงว่าข้อมูลชุดที่ 1 มีความครอบคลุมเส้นทางเป้าหมาย ดังนั้นกระบวนการหาค่าที่ได้นั้นสำหรับเส้นทางเป้าหมาย P4 ก็จะจบลง

ตัวอย่างกำหนดให้ P1 เป็นเส้นทางเป้าหมายสำหรับการหาค่าที่ได้นั้น ดังนั้นค่าสิ่งที่เป็นเงื่อนไขในเส้นทาง P1 ประกอบด้วย $\{2T, 5T, 6T\}$ เมื่อคำนวณหาผลต่างระหว่าง distance โดยใช้ชุดข้อมูลทั้ง 3 ในตารางที่ 3 จะพบว่าไม่มีข้อมูลชุดใดให้ค่าผลต่างระหว่าง distance เท่ากับศูนย์ ดังนั้นแสดงว่ายังไม่มีข้อมูลชุดใดครอบคลุมเส้นทางเป้าหมาย จึงต้องมีการดำเนินการต่อไปโดยการไขว้เปลี่ยนและการกลายพันธุ์

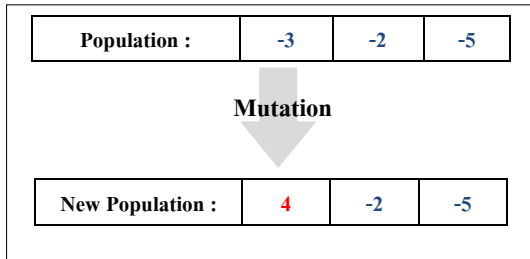
ในการไขว้เปลี่ยน จะเลือกชุดข้อมูลที่ทำให้ค่าผลต่างระหว่าง distance น้อยที่สุด 2 อันดับ และสุ่มเลือกตำแหน่งในการไขว้เปลี่ยนเพื่อให้ได้ข้อมูลชุดใหม่สำหรับการทดสอบในรอบต่อไป ตัวอย่างการไขว้เปลี่ยน แสดงดังภาพที่ 4



ภาพที่ 4 : การไขว้เปลี่ยนชุดข้อมูลแบบจุดเดียว (Single-Point)

Crossover)

กรณีที่ทำกาไรไขว้เปลี่ยนแล้ว แต่กรณีทดสอบยังไม่ครอบคลุม จะกระทำการกลายพันธุ์ โดยเลือกชุดข้อมูลที่ให้ค่าผลต่างระหว่าง distance น้อยที่สุด และสุ่มเลือกตำแหน่งในการกลายพันธุ์ ตัวอย่างการกลายพันธุ์แสดงดังภาพที่ 5



ภาพที่ 5 : การกลายพันธุ์ชุดข้อมูลแบบจุดเดียว (Single-Point Mutation)

นำชุดข้อมูลใหม่ที่ได้จากขั้นตอนที่ 4 และ 5 มาเป็นชุดข้อมูลใหม่สำหรับการหากรณีทดสอบในรอบถัดไป

กระบวนการค้นหาชุดข้อมูลสำหรับเส้นทางเป้าหมายจะถูกระทำซ้ำๆ การทำงานจะจบลงเมื่อสามารถหาชุดข้อมูลสำหรับใช้เป็นกรณีทดสอบสำหรับทุกๆ เส้นทางเป้าหมายตารางที่ 4 แสดงกรณีทดสอบที่ครอบคลุมเส้นทางเป้าหมายทั้งหมดสำหรับการทดสอบโปรแกรม Triangle

ตารางที่ 4 : กรณีทดสอบที่ครอบคลุมเส้นทางเป้าหมายสำหรับการ

Target Path	Test Case	Input values			No. of GA
		a	b	c	
P1	1	2	2	2	5
P2	2	2	4	5	6
P3	3	4	4	2	7
P4	4	-3	-2	-5	7

ทดสอบโปรแกรม Triangle

ตารางที่ 4 แสดงกรณีทดสอบที่ครอบคลุมเส้นทางทั้งหมดสำหรับการทดสอบโปรแกรม Triangle มีทั้งหมด 4 กรณีทดสอบ จำนวนรอบการดำเนินการขั้นตอนวิธีเชิงพันธุกรรม (No. of GA) เพื่อค้นหากรณีทดสอบที่ครอบคลุมเส้นทางเป้าหมาย P1 ถึง P4 มีจำนวน 5, 6, 7 และ 7 รอบตามลำดับ สำหรับเส้นทางเป้าหมาย P5 ไม่สามารถค้นหา

กรณีทดสอบที่ครอบคลุมได้ โดยใช้จำนวนรอบในการค้นหา 20 รอบซึ่งเป็นจำนวนรอบสูงสุดที่กำหนดไว้ จากการวิเคราะห์เส้นทาง P5 พบว่าเป็นเส้นทางที่เป็นไปไม่ได้ (infeasible path) ซึ่งเป็นเส้นทางที่ไม่มีกรณีทดสอบใดๆที่สามารถผ่านเส้นทางนี้ได้ ดังนั้นจึงสอดคล้องกับผลที่ได้รับจากขั้นตอนที่ได้นำเสนอ

5. สรุปผลและงานในอนาคต

บทความนี้เสนอการสร้างกรณีทดสอบโดยใช้ขั้นตอนวิธีเชิงพันธุกรรมเพื่อคัดเลือกกรณีทดสอบที่เหมาะสมสำหรับการทดสอบซอฟต์แวร์ การดำเนินการเริ่มจากนารหัสต้นทางของโปรแกรมที่ต้องการทดสอบมาแปลงเป็นกราฟควบคุมการไหล ต่อจากนั้นจะค้นหากรณีทดสอบที่ครอบคลุมทุกเส้นทางการทำงานของโปรแกรมโดยใช้ขั้นตอนวิธีเชิงพันธุกรรม เมื่อนำวิธีการที่ได้นำเสนอไปประยุกต์ใช้กับกรณีศึกษาพบว่า ขั้นตอนวิธีที่นำเสนอสามารถสร้างกรณีทดสอบสำหรับการทดสอบซอฟต์แวร์ได้อย่างมีประสิทธิภาพ

ในอนาคตได้วางแผนที่จะพัฒนาเครื่องมือเพื่อสร้างกรณีทดสอบจากวิธีการที่นำเสนอและนำไปประยุกต์ใช้กับกรณีศึกษาอื่นๆ เพื่อประเมินค่าความถูกต้องของกรณีทดสอบและพัฒนาวิธีการจัดการกับการหากรณีทดสอบสำหรับเส้นทางที่เป็นไปไม่ได้

6. เอกสารอ้างอิง

- [1] C. Ghezzi, M. Jazayeri, and D. Mandrioli. "Fundamentals of Software Engineering 2nd," Prentice Hall PTR Upper Saddle River, NJ, USA, 2002.
- [2] G. Winter, J. Periaux, M.Galan and P.Cuesta. "Genetic Algorithms in Engineering and Computer Science," Wiley Publisher, 1996.
- [3] P. Ammann and J. Offutt. "Introduction to Software Testing," Cambridge University Press, 2008.
- [4] T. J. McCabe. "A Complexity Measure," IEEE Transactions on Software Engineering, pp. 308-320, 1976.
- [5] B. Korel. "Automated Test Data Generation," IEEE Transactions on Software Engineering, pp.870-879, 1990.

การประยุกต์ใช้คิวแกรมอัลกอริทึมและโครงสร้างข้อมูลแบบต้นไม้สเปล์ย์เพื่อเพิ่ม
ประสิทธิภาพของพจนานุกรมอิเล็กทรอนิกส์ศัพท์ทางการแพทย์แผนจีน
**The Application Q-Gram Algorithm and Splay-Tree Data Structure to
Enhance Electronic Dictionary of TCM Terminology**

พีระพัฒน์ ปันดิซอน (Peerapad Pandison)¹ วรเมธ วิฑูรย์เวชณ์ (Woramet Withunwet)²

อำไพ ชอบเดียน (Aumpai Chobdean)³ และ เอกพงษ์ เชื้อสุวรรณ Ekapong Chuasuwan⁴

^{1,2,4}ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ วิทยาลัยเชียงราย

³ภาควิชาแพทย์แผนจีน คณะแพทยตะวันออก วิทยาลัยเชียงราย

¹shisurusan@hotmail.com, ²woramet_xeemlis@hotmail.com, ³hikaru_bb@hotmail.com,
⁴ekapong.chuasuwana@gmail.com

บทคัดย่อ

บทความนี้ได้นำเสนอการประยุกต์ใช้คิวแกรมอัลกอริทึมและรูปแบบการทำงานของโครงสร้างข้อมูลแบบต้นไม้สเปล์ย์เพื่อเพิ่มประสิทธิภาพของพจนานุกรมอิเล็กทรอนิกส์ศัพท์ทางการแพทย์แผนจีน โดยมีวัตถุประสงค์หลักคือการเพิ่มความเร็วในการค้นหาตามพฤติกรรมของผู้ใช้และแนะนำคำศัพท์ที่ใกล้เคียงกับความต้องการ จากการศึกษาพฤติกรรมของผู้ใช้พจนานุกรมพบว่ามักจะหาคำศัพท์ที่ซ้ำกันหรือคล้ายกันในช่วงระยะเวลาหนึ่ง ซึ่งจากการนำโครงสร้างข้อมูลแบบต้นไม้สเปล์ย์มาใช้ในการสร้างฐานข้อมูลคำศัพท์ทำให้คำศัพท์ที่ถูกค้นหาลดลง ชี้นมาอยู่อันดับต้นๆ ของโครงสร้าง ส่งผลให้ระยะเวลาในการหาคำศัพท์เร็วขึ้น และด้วยลักษณะโครงสร้างข้อมูลแบบต้นไม้ที่นำมาใช้งาน จึงทำให้การค้นหาคำศัพท์ที่ไม่ถูกค้นหานั้นยังสามารถหาได้เร็วขึ้น สำหรับการหาคำศัพท์ที่ใกล้เคียงนั้น วิธีที่นำเสนอนี้ได้ทำการเปลี่ยนคำศัพท์ภาษาจีนให้อยู่ในรูปของภาษาพินอิน ก่อนนำไปประมวลผลกับคิวแกรมอัลกอริทึมในการหาคำที่มีลักษณะใกล้เคียงกัน ผลการทดลองที่ได้จากงานนี้แสดงให้เห็นว่าวิธีที่นำเสนอสามารถช่วยลดเวลาในการค้นหาคำศัพท์จากวิธีโดยทั่วไปและยังสามารถช่วยหาคำศัพท์ที่ใกล้เคียงกับคำศัพท์ที่ผู้ใช้ต้องการเป็นอย่างดี

คำสำคัญ: คิวแกรมอัลกอริทึม, ต้นไม้สเปล์ย์, คำศัพท์แพทย์แผนจีน, พจนานุกรมอิเล็กทรอนิกส์

Abstract

This paper presented The Application Q-Gram Algorithm and Splay-Tree Data Structure to Enhancing Efficiency of Electronic Dictionary of TCM Terminology. The main purpose is to speed up to search by behavior of users and suggest the terminology that are similar to demand. The studied behavior of user dictionaries, found that they will always find duplicates or similar the terminology during the period. Which the application of splay tree data structures to the adapted terminology database, make terminology was searched frequently can up to the top of the structure. Such approaches resulting of time to determine the terminology is faster. Tree data structure implemented making search the terminology that are not searched frequently it can find it quickly. For finding that similar terminology. The method presented has changed the Chinese terminology to in the Pinyin form. Before the processing with Q-Gram algorithm for finding words that are similar. The experimental results of proposed method can reduce the searching time of the universal terminology and can also help find a terminology that is similar to terminology the user desires as well.

Keyword: Q-Gram Algorithm, Splay Tree, TCM Terminology, Electronic Dictionary.

1. บทนำ

การแพทย์แผนจีนเป็นศาสตร์ที่น่าสนใจในปัจจุบัน เพราะการแพทย์แผนจีนเป็นภูมิปัญญาที่มีการถ่ายทอดสืบต่อกันมาเป็นเวลานานกว่าหลายปี และการแพทย์แผนจีนยังได้รับการยอมรับอย่างแพร่หลายรวมทั้งประเทศไทย โดยการแพทย์แผนจีนจะมีหลักการวินิจฉัยโรคและแนวทางการรักษาโรคที่แตกต่างจากการแพทย์แผนปัจจุบัน เช่น การรักษาแบบการฝังเข็ม ยาจีน การนวดจีน และการแพทย์แผนจีนก็มีแนวคิดหลักการดูแลสุขภาพและยาจีนหลายอย่างที่มีความคล้ายคลึงกับยาไทย แต่ด้วยปัญหาด้านภาษา ความลึกซึ้งของการแพทย์แผนจีน และคำศัพท์ที่มีอยู่มากมาย ซึ่งคำศัพท์เหล่านี้จะเป็นคำศัพท์ที่ใช้บ่อย เช่น คำศัพท์ที่ใช้ในการวินิจฉัยโรคต่างๆ คำศัพท์ที่เกี่ยวกับการฝังเข็ม ฯลฯ ของการแพทย์แผนจีน จึงทำให้มีความเข้าใจยาก มีความยุ่งยากในการค้นหาคำศัพท์ และยังเสียเวลาในการเปิดหนังสือพจนานุกรมเพื่อค้นหาคำศัพท์แพทย์แผนจีน

พจนานุกรมอิเล็กทรอนิกส์เป็นเทคโนโลยีที่ได้มีการพัฒนาอย่างต่อเนื่องเพื่ออำนวยความสะดวกแก่ผู้ที่ต้องการหาความหมายของคำศัพท์ต่างๆ ซึ่งรูปแบบการทำงานของการค้นหาข้อมูลคำศัพท์ของพจนานุกรมอิเล็กทรอนิกส์โดยทั่วไปจะมีรูปแบบความเร็วในการทำงานเท่ากับ $O(n)$ กล่าวคือเมื่อต้องการค้นหาคำศัพท์ในฐานข้อมูล ระบบจะทำการเช็คข้อมูลคำศัพท์ในฐานข้อมูลที่ละระเบียบข้อมูล ซึ่งเมื่อฐานข้อมูลคำศัพท์มีข้อมูลในปริมาณที่มาก ความเร็วในการทำงานของระบบก็จะช้าลง และปัญหาอีกอย่างหนึ่งที่สำคัญของพจนานุกรมทั่วไปคือ ไม่สามารถที่จะหาคำศัพท์ที่มีความหมายใกล้เคียงกับคำศัพท์ที่ผู้ใช้ต้องการ

จากผลการทดสอบสังเกตลักษณะการค้นหาคำศัพท์จากกลุ่มผู้ใช้ของบุคลากรและนักศึกษาแพทย์แผนจีน คณะแพทยศาสตร์วชิรพยาบาล มหาวิทยาลัยนวมินทราธิราช พบว่าในทุกๆ ช่วงเวลาหนึ่งของการใช้งาน กลุ่มผู้ใช้จะมีการค้นหาคำศัพท์ที่เหมือนหรือคล้ายกันเป็นจำนวนมาก หรืออาจจะกล่าวได้ว่าคำศัพท์บางกลุ่มมีความถี่ในการค้นหาที่สูง

ในงานวิจัยนี้จึงได้มีแนวคิดประยุกต์ใช้โครงสร้างข้อมูลแบบต้นไม้สเปย์เข้ามาช่วยในการค้นหาคำศัพท์เพื่อทำให้ความเร็วในการค้นหาคำศัพท์เร็วขึ้น และประยุกต์ใช้คิวแกรม

อัลกอริทึมในการช่วยค้นหาคำศัพท์ที่ใกล้เคียงกับคำศัพท์ที่ผู้ใช้ต้องการหรือคำศัพท์ที่ใกล้เคียงกับคำศัพท์ที่ผู้ใช้อาจพิมพ์ผิด

2. วรรณกรรมที่เกี่ยวข้อง

2.1 โครงสร้างข้อมูลแบบต้นไม้สเปย์ (Splay Tree Data Structure) [1]

โครงสร้างข้อมูลแบบต้นไม้สเปย์ เป็นต้นไม้ค้นหาแบบทวิภาคที่มีลักษณะโครงสร้างปรับปรุงตัวเอง (Self-Adjusting Tree) คือสามารถปรับโครงสร้างทุกครั้งที่มีการเรียกใช้งาน ไม่ว่าจะเป็นการเพิ่ม การลบ รวมถึงการค้นหา โดยจะทำการนำโหนด (Node) ที่ถูกเรียกใช้งานขึ้นมาเป็นโหนดราก (Root Node) นอกจากนั้นตัวโครงสร้างของต้นไม้เมื่อถูกเรียกใช้งานบ่อยๆ ต้นไม้จะค่อยปรับรูปร่างไม้ที่มีลักษณะเตี้ยและแผ่ขยายออก

จุดเด่นของต้นไม้สเปย์คือการปรับให้โหนดที่เพิ่งถูกใช้บริการขึ้นมาเป็นโหนดราก ซึ่งทำให้การค้นหาข้อมูลที่เพิ่งค้นหาไปใหม่นั้นรวดเร็วขึ้น จากแนวคิดของการค้นหาแบบโดยทั่วไปที่ว่า กำหนดให้ความถี่ของโหนดที่ต้องการค้นหาเท่ากับกันหมด แต่โหนดที่ถูกค้นหาเยอะๆ ก็จะมีโอกาสที่ถูกค้นพบง่ายขึ้น ดังนั้นการทำให้ข้อมูลที่ถูกค้นบ่อยๆ ขึ้นไปอยู่ตำแหน่งบนๆ ของโครงสร้างต้นไม้ ก็จะทำให้การค้นหาในอนาคตรวดเร็วมากขึ้น ซึ่งความเร็วในการค้นหาของต้นไม้สเปย์มีความเร็วเฉลี่ยเท่ากับ $O(\log n)$

2.2 คิวแกรมอัลกอริทึม (Q-Gram Algorithm) [2-4]

คิวแกรมอัลกอริทึม เป็นการเรียงลำดับตัวอักษร ของข้อความ คือ ตัวอักษรที่เป็นตัวขึ้นต้นฉบับนำมาเรียงลำดับเป็นตัวอักษรย่อย โดยใช้วิธีการคำนวณ edit distance ซึ่งสามารถคำนวณได้โดยภาษา SQL โดยมีหลักการคำนวณดังขั้นตอนต่อไปนี้ การเชื่อมโยงข้อมูลแบบกำหนดตายตัว การทดสอบว่า $(a, b) \in M$ หรือไม่นั้น จะใช้กฎเอกลักษณ์หรือตัวบ่งชี้ที่ไม่มีค่าซ้ำ (Unique Identifier) เช่น เลขประจำตัวประชาชน โดยกำหนดเงื่อนไขว่า ตัวบ่งชี้ใน a และ b ต้องมีค่าตรงกัน (Exact Match) จึงจะถือว่าทั้งสองระเบียบเป็นข้อมูลของคนคนเดียวกัน

แต่ในทางปฏิบัติ ข้อมูลจากแหล่งต่างๆ อาจมีความคลาดเคลื่อนไปบ้าง อาทิเช่น

กรณีที่มีการสะกดเพี้ยน เช่น “พีระพัฒน์ ปันดิสอน” ในระเบียบ a และ “พีระพัฒน์ ปันดิสอน” ในระเบียบ b

กรณีที่ตัวบ่งชี้หนึ่งเป็นเซตย่อยของอีกตัวบ่งชี้หนึ่ง เช่น “พีระพัฒน์ ปันดิสอน” ในระเบียบ a และ “ดร.พีระพัฒน์ ปันดิสอน” ในระเบียบ b

กรณีของการใช้ชื่อเต็ม ชื่อย่อ และลำดับของชื่อ-นามสกุล ในภาษาอังกฤษแตกต่างกัน เช่น “James A. Smith”, “James Smith”, “J. A. Smith” และ “Smith, J. A.” เป็นต้น

การเชื่อมโยงข้อมูลแบบเชิงความน่าจะเป็นอาศัยการจับคู่ตัวบ่งชี้แบบประมาณ (Approximate Match) โดยมีหลักการ ดังนี้

ให้ตัวบ่งชี้ที่ใช้เชื่อมโยงข้อมูล a และ b มีอยู่ k ตัว

ความต่างกันของตัวบ่งชี้ ตัวที่ i คือ Δi ความต่างกันของตัวบ่งชี้ทั้ง k ตัว จึงเท่ากับเวกเตอร์ $(\Delta 1, \dots, \Delta k)$ ซึ่งเวกเตอร์ดังกล่าว จะถูกประมวลผลด้วยฟังก์ชันการตัดสินใจ เพื่อหาข้อสรุปว่า a และ b น่าจะเป็นข้อมูลของคนคนเดียวกันหรือไม่

2.3 ลักษณะคำศัพท์การแพทย์แผนจีน [5]

อักษรจีนเป็นตัวอักษรที่มาจากภาพวาดและวิวัฒนาการเรื่อยมาจนเป็นตัวอักษรจีนในปัจจุบัน โดยแต่ละขีด แต่ละเส้นของตัวอักษรจีนไม่สามารถนำมาผสมกันแล้วออกเสียงเป็นภาษาจีนได้ ด้วยเหตุนี้จึงจำเป็นต้องเรียนรู้ระบบการออกเสียงตัวอักษรภาษาจีน (พินอิน 拼音) ซึ่งได้ดัดแปลงมาจาก International Phonetic Alphabets เพื่อช่วยในการอ่านออกเสียงตัวอักษรจีนนั้น ๆ

พยัญชนะต้นของภาษาจีนมี 21 เสียง

b (บิว / ปอ) p (พั่ว / ฟอ) m (มั่ว / มอ) f (ฟิว / ฟอ) d (เตอ) t (เทอ) n (เนอ) l (เลอ) g (เกอ) k (เคอ) h (เฮอ) j (จี) q (ชี) x (ซี) z (จื่อ) c (ชือ) s (ซือ) zh (จื่อ) ch (ชือ) sh (ชือ) r (ยื่อ)

พยัญชนะต้นกึ่งสระมี 2 เสียง คือ y (อี อี) และ w (อู อู)

สระ แบ่งออกเป็นสระเดี่ยว และสระผสม ดังนี้

单韵母 dān yùnmǔ ดานอวี่นหมู่ สระเดี่ยว มี

6 เสียง ดังนี้ a (อา) o (โอ) e (เออ) i (อี) u (อู) ü (อวี)

双韵母 shāung yùnmǔ ชวงอวี่นหมู่

สระผสม สังเกตได้ว่าการออกเสียงสระผสม จะเกิดจากเทคนิคการลากเสียงของสระแต่ละเสียงมาชนกัน

a (อา) + o (โอ) = ao (อาว) i (อี) + u (อู) = iu (อิว) e (เออ) + i (อี) = ei (เอย)

a (อา) + i (อี) = ai (ไอย) u (อู) + i (อี) = ui (อุย) o (โอ) + u (อู) = ou (โอว)

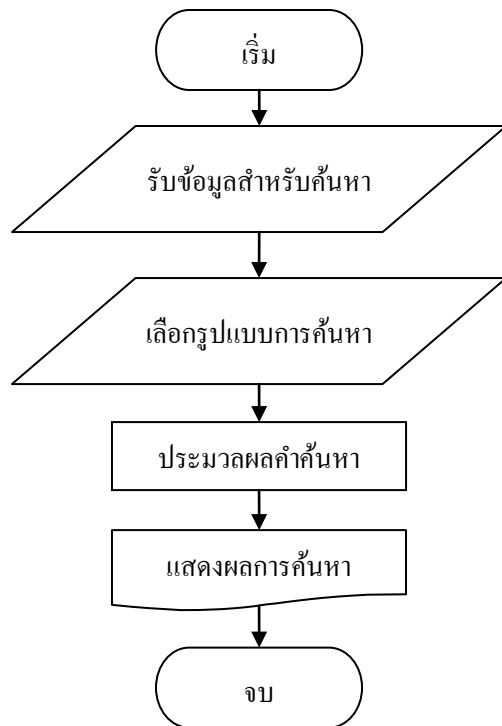
u (อู) + o (โอ) = uo (อัว) u (อู) + a (อา) = ua (อวา) เป็นต้น

3. ขั้นตอนการออกแบบและพัฒนาระบบ

การทำงานของพจนานุกรมอิเล็กทรอนิกส์ศัพท์ทางการแพทย์แผนจีนในงานที่นำเสนอ มีรูปแบบการทำงานดังภาพที่ 1 ที่แสดงภาพรวมทั้งหมดของการทำงานของพจนานุกรมอิเล็กทรอนิกส์ศัพท์ทางการแพทย์แผนจีน โดยได้ออกแบบการทำงานเริ่มจากการป้อนคำศัพท์ให้ระบบทำการรับข้อมูลสำหรับนำไปค้นหาในฐานข้อมูล จากนั้นผู้ใช้จะทำการเลือกรูปแบบการค้นหาตามความต้องการ ซึ่งในงานนี้ได้กำหนดรูปแบบการค้นหาออกเป็น 4 รูปแบบ ได้แก่

- ค้นหาจากคำศัพท์ภาษาจีน
- ค้นหาจากคำอ่านภาษาไทย
- ค้นหาจากความหมายภาษาไทย
- ค้นหาจากความหมายภาษาอังกฤษ

เมื่อทำการเลือกรูปแบบการค้นหา ระบบจะทำการประมวลผลคำค้นหา โดยมีวิธีการค้นหาในฐานข้อมูล 2 แบบคือแบบ แบบที่ 1 ค้นหาโดยใช้โครงสร้างข้อมูลต้นไม้สเปล์ช และแบบที่ 2 ค้นหาโดยใช้คิวแกรมอัลกอริทึมเพื่อหาคำศัพท์อื่นๆ ที่ใกล้เคียงกับที่ผู้ใช้ต้องการ แล้วจึงทำการแสดงผลลัพธ์ให้ผู้ใช้



ภาพที่ 1: ลำดับขั้นตอนการทำงานของพจนานุกรมอิเล็กทรอนิกส์ศัพท์ทางการแพทย์แผนจีน

3.1 ภาษาคอมพิวเตอร์และระบบฐานข้อมูลที่ใช้ใน

การพัฒนาาระบบ

ในงานที่นำเสนอนี้มีรูปแบบการทำงานในลักษณะของเว็บแอปพลิเคชัน เพื่อเอื้ออำนวยความสะดวกให้กับผู้ใช้ ผู้พัฒนาจึงได้เลือกใช้ภาษา PHP (Personal Home Page) ซึ่งเป็นภาษาคอมพิวเตอร์แบบ Open Source ที่ได้รับความนิยมอย่างแพร่หลาย [6]

ทางด้านฐานข้อมูลในงานนี้ได้เลือกใช้ฐานข้อมูลของ MySQL ซึ่งเป็นระบบจัดการฐานข้อมูลเชิงสัมพันธ์ที่ได้รับความนิยมในการนำมาพัฒนาร่วมกับเว็บแอปพลิเคชันเช่นกัน MySQL ใช้ภาษา SQL ในการจัดการชุดคำสั่งต่างๆ โดยเป็นระบบฐานข้อมูลแบบ Open Source ที่อยู่ภายใต้ลิขสิทธิ์ของบริษัทซันไมโครซิสเต็มส์ (Sun Microsystems, Inc.) [7]

3.2 คิวแกรมอัลกอริทึม

ในงานนี้ได้ดัดแปลงประยุกต์ใช้คิวแกรมอัลกอริทึมสร้างเป็นฟังก์ชันเพื่อช่วยคำนวณหาจำนวนข้อมูลย่อยที่คล้ายกันของคำศัพท์ที่ผู้ใช้ป้อนเข้าสู่ระบบเทียบกับคำศัพท์ที่มีในฐานข้อมูล โดยมีรูปแบบของอัลกอริทึมดังแสดงในภาพที่ 2

```

SearchPosQGram (keyword  $k$ , TextCompare  $t$ ,
NumberCharQgram  $q$ )
1: Create variable of length of  $k$ 
2: Add character “##” to the front and the back in  $k$ 
3: Loop  $i = 0$  to number of length of  $k + q - 1$ 
4: Create array variable “SubText” for keep string
from sub string  $k$ , start character position  $i$ 
amount  $q$ 
5: End loop
6: Create Q-Gram position variable “ $np$ ”
7: Loop  $i = 0$  to array number of SubText
8: Add character “##” to the front and the back in  $t$ 
9: If can find SubText[ $i$ ] in  $t$  than  $np++$ 
10: End loop
11: Return  $np$ 
  
```

ภาพที่ 2: อัลกอริทึมการหาจำนวนข้อมูลย่อยที่คล้ายกันของคำศัพท์ที่ผู้ใช้ป้อนเข้าสู่ระบบโดยประยุกต์จากคิวแกรมอัลกอริทึม

จากภาพที่ 2 เป็นอัลกอริทึมสำหรับหาจำนวนตำแหน่งของคำศัพท์ในฐานข้อมูลที่คล้ายกับคำศัพท์ที่ผู้ใช้ต้องการค้นหา โดยการทำงานของอัลกอริทึมจะทำการเพิ่มอักขระ “##” ในส่วนของข้างหน้าและข้างหลังข้อมูลคำศัพท์ที่ต้องการค้นหา (k) จากนั้นจะทำการสร้างตัวแปรประเภทอาร์เรย์เพื่อเก็บชุดข้อมูลของตัวอักษรที่ถูกตัดออกเป็นช่วงๆ โดยมีความยาวของการตัดอักษรเท่ากับ q ที่ผู้ใช้กำหนด ในงานนี้ได้กำหนดค่า $q = 2 - 5$ ตามค่าพื้นฐานของคิวแกรมอัลกอริทึม [2]

จากนั้นจะทำการเปรียบเทียบข้อมูลชุดข้อมูลย่อยของตัวอักษรที่ตัดมานี้กับคำศัพท์แต่ละตัวในฐานข้อมูล เพื่อหาความคล้ายกันของคำศัพท์ ซึ่งเมื่อเปรียบเทียบครบทั้งฐานข้อมูล อัลกอริทึมจะทำการคืนค่าจำนวนตำแหน่งของชุดอักษรย่อยที่คล้ายกันของแต่ละคำศัพท์กลับมาให้

การคัดเลือกคำศัพท์ที่ใกล้เคียง หลังจากที่ได้ค่าจำนวนคำศัพท์ที่ใกล้เคียงของแต่ละคำศัพท์แล้ว จะทำการจัดเรียงข้อมูลคำศัพท์ตามค่าที่ได้จากมากไปหาน้อยเพื่อคัดเลือกคำศัพท์ที่ใกล้เคียงมาแสดงผลให้ผู้ใช้ โดยการคัดเลือกคำศัพท์ที่คล้ายนี้ ได้ใช้การกำหนดค่าคัดเลือก (threshold of selected) = 5 กล่าวคือหากให้ค่าที่กำหนดการคัดเลือกมามีค่ามาก จำนวนคำศัพท์ที่คล้ายคลึงจะมีมาก ซึ่งบางคำศัพท์ที่ตัดเลือกมาเป็นคำศัพท์ที่ไม่ได้มีความหมายใกล้เคียงกับความต้องการของผู้ใช้ ซึ่งหากพิจารณาจากค่าที่กำหนดในงานนี้ จากค่า q ที่เท่ากับ 3 และค่าคัดเลือก = 5 ทำให้มีโอกาสความเป็นไปได้ที่ความต่างของตัวอักษรของคำศัพท์จะต่างกันได้ถึง 15 ตัวอักษร (1

ตำแหน่งของชุดข้อมูลย่อยเท่ากับ 3 ตัวอักษร) ซึ่งหากคำศัพท์ใด ๆ มีความต่างกันถึง 15 ตัวอักษรแล้ว จึงไม่น่าจะเป็นคำศัพท์ที่มีความหมายหรืออยู่ในกลุ่มเดียวกับความต้องการของผู้ใช้

3.3 อัลกอริทึมในการปรับโครงสร้างฐานข้อมูล

SetFirst (LastTerminology *LT*)

- 1: Get position and cumulative frequency of *LT*
- 2: Change *LT* Data to first record
- 3: Return cumulative frequency of *LT*

SortData (CumulativeFrequency *cf*)

- 1: If amount of record in database equal zero
- 2: Return order of sort data
- 3: Else
- 4: Divide into two parts in database, by number of cumulative frequency of terminology are greater than and less than *cf*.
- 5: Ascending sort data of cumulative frequency is greater.
- 6: Get first record in data of cumulative frequency is greater after sorting. (*hr*)
- 7: Descending sort data of cumulative frequency is less.
- 8: Get first record in data of cumulative frequency is less after sorting. (*lr*)
- 9: Set *hr* and *lr* to order of sort data.
- 10: Get cumulative frequency of *hr* (*cfh*) and *lr* (*clf*)
- 11: Delete record *hr* and *lr*
- 12: SortData(*cfh*)
- 13: SortData(*clf*)

ภาพที่ 3: อัลกอริทึมของการปรับลำดับข้อมูลในฐานข้อมูล

ในการเพิ่มความเร็วในการค้นหาคำศัพท์ในงานนี้ได้คิดแปลงประยุกต์วิธีการของโครงสร้างข้อมูลแบบต้นไม้สเปลิย์มาทำการปรับโครงสร้างของฐานข้อมูล โดยมีรูปแบบการทำงานคือ ทุกครั้งที่มีการค้นหาคำศัพท์ใด ๆ จะมีการเก็บจำนวนการค้นหาคำศัพท์นั้น ๆ ไว้ เพื่อนำมาจัดเรียงลำดับของข้อมูลคำศัพท์ในฐานข้อมูล จากนั้นเมื่อมีการค้นหาคำศัพท์คำศัพท์ที่ถูกค้นหาล่าสุดจะถูกปรับให้ขึ้นมายู่ลำดับที่ 1 ของฐานข้อมูล และข้อมูลที่เหลือ จะถูกปรับการจัดเรียงของลำดับใหม่ โดยมีอัลกอริทึมในการทำงานดังแสดงในภาพที่ 3

อัลกอริทึมในการจัดเรียงข้อมูลในฐานข้อมูลหลังจากที่มีการค้นหา โดยหลังจากที่มีการปรับให้ข้อมูลที่ถูกลบออกแล้ว อัลกอริทึมจะทำการแบ่งข้อมูลออกเป็นสองชุด คือชุดข้อมูลที่คำศัพท์มีความถี่สะสมน้อยกว่าคำศัพท์ที่เพิ่งถูกค้นหาไปและชุดข้อมูลที่คำศัพท์มีความถี่สะสมมากกว่า จากนั้นจะทำการจัดเรียงข้อมูลทั้งสอง

เพื่อคัดเลือกเอาข้อมูลอันดับที่หนึ่งของทั้งสองเก็บไว้ในรายการจัดเรียง การทำงานของอัลกอริทึมจะเป็นแบบรีเคอร์ซีฟ (Recursive)

3.4 ข้อมูลสำหรับการทดลอง

ในงานนี้ได้ออกแบบฐานข้อมูลในการเก็บรวบรวมคำศัพท์การแพทย์แผนจีนจากหนังสือ “พจนานุกรมศัพท์การแพทย์แผนจีน (จีน-ไทย-อังกฤษ)” [5] โดยมีจำนวนข้อมูลคำศัพท์ทั้งหมด 1,015 คำ ซึ่งในฐานข้อมูลจะประกอบไปด้วยคำศัพท์ภาษาจีน, คำอ่านพินอิน, คำอ่านภาษาไทย, ความหมายภาษาไทย และความหมายภาษาอังกฤษ

4. ผลการทดลอง

ในงานที่นำเสนอนี้ได้ออกแบบการทดลองเพื่อทดสอบประสิทธิภาพของระบบออกเป็น 2 ประเด็น คือ 1) จำนวนคำศัพท์ที่ใกล้เคียงที่ได้ และ 2) ความเร็วที่ใช้ในการค้นหา โดยทำการทดสอบบนเครื่องคอมพิวเตอร์ที่ใช้ระบบปฏิบัติการ Microsoft Windows 7 ประเภท 64 Bit มีหน่วยความจำ 16 GB และหน่วยประมวลผลหลัก 2.10 GHz

4.1 การทดลองเพื่อหาจำนวนคำศัพท์ที่ใกล้เคียง

เพื่อให้ได้ความเหมาะสมของจำนวนคำศัพท์ที่ใกล้เคียงตามความต้องการของผู้ใช้ จึงได้ทดสอบโดยการกำหนดค่า q ตั้งแต่ 2 – 5 ซึ่งผลลัพธ์ที่ได้ในการทดลอง ดังแสดงในตารางที่ 1

ตารางที่ 1: ตารางแสดงจำนวนคำใกล้เคียงตามค่า q ที่แตกต่าง

ค่า q	จำนวนคำศัพท์ที่คล้ายน้อยที่สุด	จำนวนคำศัพท์ที่คล้ายมากที่สุด	จำนวนเฉลี่ยคำศัพท์ที่คล้าย
2	11	100	50
3	3	30	8
4	1	15	4
5	1	12	2

ในตารางที่ 1 แสดงจำนวนคำศัพท์ที่ใกล้เคียงกับคำที่ผู้ใช้ต้องการ ซึ่งแสดงให้เห็นว่าเมื่อกำหนดค่า q ในคิวแกรม อัลกอริทึมมีค่ามากขึ้นทำให้ได้จำนวนคำศัพท์น้อยลง แต่หากกำหนดค่า q น้อย ก็จะทำให้ได้จำนวนคำศัพท์ที่มาก ซึ่งเป็นจำนวนที่ไม่เหมาะสมกับการนำมาแสดงผลให้ผู้ใช้ ในการ

ทดลองนี้จึงได้เลือกใช้ค่า q เท่ากับ 3 ซึ่งได้จำนวนเฉลี่ยของคำศัพท์ที่เจอจำนวน 8 คำศัพท์

4.2 การทดลองเพื่อทดสอบความเร็วที่ใช้ในการค้นหา

ในการทดลองเพื่อทดสอบประสิทธิภาพด้านความเร็วของการค้นหาของวิธีที่นำเสนอนี้ ได้ทำการทดสอบโดยการแบ่งข้อมูลสำหรับการทดสอบออกเป็น 5 ชุด เพื่อหาความเร็วเฉลี่ยของการค้นหา โดยแสดงผลลัพธ์ในตารางที่ 2

ตารางที่ 2: ตารางแสดงเวลาเฉลี่ยในการค้นหาคำศัพท์ เปรียบเทียบกับวิธีการค้นหาแบบธรรมดาที่วิธีที่นำเสนอ

ชุดข้อมูลที่	จำนวนข้อมูล (Record)	เวลาที่ใช้ในการค้นหา (Sec)	
		แบบธรรมดา	วิธีที่นำเสนอ
1	200	0.00531	0.00120
2	400	0.00832	0.00173
3	600	0.01932	0.00331
4	800	0.03714	0.00351
5	1,015	0.05124	0.00371

ในตารางที่ 2 แสดงให้เห็นว่าวิธีที่นำเสนอนี้ สามารถช่วยลดระยะเวลาในการค้นหาคำศัพท์ โดยเมื่อเปรียบเทียบกับการค้นหาแบบธรรมดาที่จำนวนข้อมูล 1,015 record โดยวิธีธรรมดาในที่นี้คือใช้การค้นหาตามปกติของซอฟต์แวร์ฐานข้อมูล MySQL ผลลัพธ์ที่ได้จะเห็นได้ว่าวิธีที่ได้นำเสนอนี้ใช้เวลาเฉลี่ยในการค้นหาคำศัพท์ในฐานข้อมูลที่ได้นำมาทดสอบเพียง 0.00371 วินาที

5. สรุปผลการทดลอง

ในบทความนี้ได้นำเสนอวิธีเพิ่มประสิทธิภาพให้กับพจนานุกรมอิเล็กทรอนิกส์คำศัพท์ทางด้านแพทย์แผนจีน โดยการประยุกต์ใช้คิวแกรมอัลกอริทึมช่วยในการหาคำศัพท์ที่ใกล้เคียงกับคำศัพท์ที่ผู้ใช้ต้องการ และประยุกต์ใช้การทำงานของโครงสร้างข้อมูลแบบต้นไม้สเปล์ช่วยในการลดระยะเวลาในการค้นหาคำศัพท์

จากผลการทดลองแสดงให้เห็นว่าวิธีที่ได้นำเสนอสามารถช่วยในการค้นหาคำศัพท์ได้เป็นอย่างดี โดยสามารถช่วยหาจำนวนคำศัพท์ที่ใกล้เคียงกับที่ผู้ใช้ต้องการ โดยจากการทดลองเมื่อกำหนดค่า $q = 3$ จะได้จำนวนเฉลี่ยของคำศัพท์ที่ใกล้เคียง

กับความต้องการผู้ใช้ 8 คำศัพท์ และสามารถลดความเร็วในการค้นหา โดยใช้ความเร็วเฉลี่ยในการค้นหาข้อมูลพื้นฐานข้อมูลเพียง 0.00371 วินาที

เอกสารอ้างอิง

- [1] Weiss, Mark Allen, and Susan Hartman. *Data structures and problem solving using Java*. Addison-Wesley, 2006.
- [2] Burkhardt, Stefan, et al. "q-gram based database searching using a suffix array (QUASAR)." *Proceedings of the third annual international conference on Computational molecular biology*. ACM, 1999.
- [3] Sutinen, Erkki, and Jorma Tarhio. "On using q-gram locations in approximate string matching." *Algorithms—ESA'95*. Springer Berlin Heidelberg, 1995. 327-340.
- [4] Gravano, Luis, et al. "Approximate string joins in a database (almost) for free." *VLDB*. Vol. 1. 2001.
- [5] วิลาวัณย์ จิงประเสริฐ, เกา หย่งเสียง และคณะ, *พจนานุกรมศัพท์การแพทย์แผนจีน (จีน-ไทย-อังกฤษ)*, ชมชนสหกรณ์การเกษตรแห่งประเทศไทย จำกัด, กรุงเทพมหานคร, 2554
- [6] บัญชา ปะสิละเตสัง, *พัฒนาเว็บแอปพลิเคชันด้วย PHP ร่วมกับ MySQL และ Dreamweaver*, ซีเอ็ดดูเคชั่น, บมจ., 2553.
- [7] ชานูชัย สุภอรรถกร, *จัดการฐานข้อมูลด้วย MYSQL*, ชิมพลีฟาย, สนพ., 2554.

ต้นแบบเว็บที่สามารถทำงานในสถานะออฟไลน์โดยใช้เอชทีเอ็มแอล 5

A Prototype of Offline Tolerant Web using HTML5

ปาโมกษ์ รัตนศรียาภิบาล (Pamok Rattanatriyapibal)¹ และมาลีรัตน์ โสคานิล (Maleerat Sodanil)²

^{1,2}ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
pamok.r@rmutp.ac.th¹, msn@kmutnb.ac.th²

บทคัดย่อ

การทำงานของเว็บประยุกต์จะต้องมีการเชื่อมต่อกับระบบเครือข่ายตลอดเวลา เพราะมีรูปแบบการทำงานในลักษณะระบบรับ-ให้บริการ (Client-Server) ซึ่งหากอยู่ในสถานะออฟไลน์ก็จะไม่สามารถใช้งานเว็บประยุกต์นั้นได้ โดยในงานวิจัยจะทำการพัฒนาต้นแบบเว็บที่สามารถทนต่อสถานะออฟไลน์โดยใช้คุณลักษณะของเอชทีเอ็มแอล 5 ประกอบด้วย การใช้ ApplicationCaches ในการแคชไฟล์เว็บไซต์ที่ฝั่งผู้ใช้ เพื่อให้สามารถเข้าใช้งานเว็บได้ในขณะออฟไลน์ และใช้ WebSQL ในการสำรองข้อมูลเพื่อรองรับการจัดการกับข้อมูลในขณะออฟไลน์ และประเมินโดยใช้สถานการณ์จำลองที่ประกอบด้วย การ ค้นหา เพิ่ม อ่าน แก้ไข และลบข้อมูล ในขณะออฟไลน์เพื่อพิสูจน์ว่าต้นแบบเว็บสามารถทำงานพื้นฐานต่าง ๆ ในขณะออฟไลน์ได้ตามที่คาดหมายไว้ พบว่าต้นแบบเว็บสามารถทำงานพื้นฐานต่าง ๆ ได้ในขณะออฟไลน์

คำสำคัญ: เอชทีเอ็มแอล 5 ApplicationCaches WebSQL ออฟไลน์ เว็บไซต์

Abstract

The commonly web application must have network connection all time when usage due to the web application using Client-Server model. If there is no network connection or connection is offline user will cannot access that web application. In this research will develop a prototype of offline tolerant web using HTML5 containing two new features are ApplicationCaches use to cache the web into user side for support offline access and WebSQL use to store the data into user side for manipulation data when connection is offline. Evaluation

will use scenarios of the five basic functions contain search, create, read, update and delete. Test when connection is offline to prove that scenarios can be achieved. The result is all scenarios can be achieved.

Keyword: HTML5, ApplicationCaches, WebSQL, Offline, Website.

1. บทนำ

การทำงานของเว็บประยุกต์โดยปกติแล้วจะต้องมีการเชื่อมต่อกับเครือข่ายตลอดเวลาไม่ว่าจะเป็นเครือข่ายอินเทอร์เน็ต หรืออินทราเน็ต เพราะมีรูปแบบการทำงานในลักษณะระบบรับ-ให้บริการ โดยที่ระบบให้บริการจะเป็นผู้ให้ข้อมูล ประมวลผลข้อมูลรวมถึงภาษาโปรแกรมต่าง ๆ แล้วส่งไปยังผู้รับที่เข้าใช้บริการผ่านเว็บเบราว์เซอร์ โดยที่เว็บเบราว์เซอร์จะนำข้อมูลและภาษาโปรแกรมที่ได้รับมาสร้างเป็นหน้าจอเว็บประยุกต์ ด้วยรูปแบบการทำงานในลักษณะนี้ระบบสารสนเทศต่าง ๆ จึงนิยมพัฒนาให้อยู่ในรูปเว็บประยุกต์ เพราะผู้ใช้จะสามารถเข้าถึงได้ง่าย หลากหลายแพลตฟอร์ม ไม่ต้องทำการติดตั้งระบบ และหากผู้พัฒนาต้องการแก้ไขระบบก็สามารถแก้ไขที่ระบบให้บริการได้เลย

อย่างไรก็ตามการใช้งานเว็บประยุกต์นั้นก็ยังมีข้อจำกัดที่จำเป็นต้องมีการเชื่อมต่อกับเครือข่ายตลอดเวลา หากการเชื่อมต่อมีปัญหา หรืออยู่ในสถานะออฟไลน์ ไม่ว่าจะด้วยสาเหตุใดก็ตามผู้ใช้จะไม่สามารถเข้าใช้งานเว็บนั้น ๆ ต่อไปได้ โดยเฉพาะอย่างยิ่งในปัจจุบันมีการใช้การเชื่อมต่อผ่านอุปกรณ์เคลื่อนที่พกพาต่าง ๆ ที่มีการเชื่อมต่อไม่เสถียร โดยในระหว่างการใช้งานหากการเชื่อมต่อขาด ข้อมูลที่กำลังจะบันทึกก็อาจจะสูญหาย หรือข้อมูลที่เคยเปิดดูก่อนหน้านี้จะไม่สามารถดูได้

ซึ่งที่ผ่านมาได้มีการพยายามแก้ไขปัญหานี้เช่น Google Gears API [1] ที่ถูกพัฒนาโดย Google จะช่วยให้สามารถเก็บข้อมูลไว้บนเครื่องผู้ใช้เพื่อสนับสนุนการใช้งานแบบออฟไลน์

แต่ต่อมาได้ยกเลิกการพัฒนาเพราะมีเอชทีเอ็มแอล 5 มาแทนที่ [2] โดยที่เอชทีเอ็มแอล 5 นั้นจะมีคุณลักษณะใหม่ที่ถูกกำหนดเข้ามาซึ่งสามารถทำงานแทน Google Gears API โดยที่ผู้ใช้ไม่จำเป็นต้องติดตั้งโปรแกรมเสริมใด ๆ

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 เอชทีเอ็มแอล 5

มาตรฐานใหม่ของภาษาเอชทีเอ็มแอล ที่ทาง W3C ได้วางแผนการแนะนำให้เอชทีเอ็มแอล 5 ใช้เป็นมาตรฐานใหม่ในปี ค.ศ. 2014 [3] โดยมีลักษณะเหมือนมาตรฐานตัวก่อนหน้าทั้ง เอชทีเอ็มแอล 4.01 และ เอ็กซ์เอชทีเอ็มแอล 1.1 ที่ใช้ในการจัดโครงสร้างและการแสดงผลของเนื้อหาสำหรับเว็บไซต์ เอชทีเอ็มแอล 5 นั้นถูกออกแบบมาให้สามารถทำงานหลาย ๆ อย่าง โดยที่ไม่จำเป็นต้องพึ่งการติดตั้งโปรแกรมอื่น ๆ เสริมเหมือนในที่ผ่านมา นอกจากนั้นยังสามารถใช้งานบนแพลตฟอร์มทุกชนิด และยังสามารถสนับสนุนทำงานแบบออฟไลน์โดยจะมีคุณลักษณะใหม่เพิ่มเติมเข้ามาหลายด้าน

2.2 ApplicationCaches

คุณลักษณะ ApplicationCaches คือการแคชไฟล์ต่าง ๆ ไว้บนฝั่งผู้ใช้ช่วยให้สามารถเข้าใช้งานเว็บได้ในสภาวะออฟไลน์สามารถกำหนดได้ว่าจะให้ไฟล์ใดถูกบันทึก หรือไม่ถูกบันทึก โดยการกำหนดค่าบน Manifest โดยมีการแบ่งคำสั่งในการกำหนดออกเป็น 3 คำสั่งดังนี้ [4]

- CACHE MANIFEST จะเป็นการระบุรายการไฟล์ที่จะทำการบันทึกลงในฝั่งผู้ใช้โดยจะต้องระบุทุกไฟล์ที่จะให้บันทึก ไม่สามารถระบุเป็น Wildcard ได้
- NETWORK จะเป็นการระบุรายการไฟล์ที่จะต้องมีการเชื่อมต่อกับระบบเครือข่ายเท่านั้น สามารถระบุเป็น Wildcard โดยใช้เครื่องหมาย “*”
- FALLBACK จะเป็นการระบุว่าจะให้แสดงหน้าเว็บไซต์ใด เมื่อไม่มีการเชื่อมต่อกับระบบเครือข่าย เช่น ในภาพที่ 1 หากไม่มีการเชื่อมต่อ และผู้ใช้เปิดหน้า “/html/” จะแสดงผลหน้า “/offline.html” แทน

```

CACHE MANIFEST
# 2012-02-21 v1.0.0
/theme.css
/logo.gif
/main.js

NETWORK:
login.asp

FALLBACK:
/html/ /offline.html

```

ภาพที่ 1: Manifest File

2.3 WebSQL

คุณลักษณะ WebSQL ในเอชทีเอ็มแอล 5 อาจจะเรียกว่า Web Database โดยจะเป็นฐานข้อมูลที่ถูเก็บลงในเว็บเบราว์เซอร์ของผู้ใช้ สามารถรองรับข้อมูลที่มีความสัมพันธ์ (Relational Database) และรองรับคำสั่งแบบ SQL ช่วยให้ผู้พัฒนาสามารถสร้างโปรแกรมประยุกต์ที่มีความหลากหลายและประสิทธิภาพมากยิ่งขึ้น [5]

Prof...	Title	FirstNa...	La...
1	Mr.	Pamok	Ra
2	Mr.	Gustavo	Ac
3	Ms.	Cather...	Ab
4	Ms.	Kim	Ab
5	Sr.	Humb...	Ac
6	Sra.	Pilar	Ac
7	Ms.	Frances	Ac
8	Ms.	Marga...	Sm
9	Ms.	Carla	Ac
10	Mr.	Jay	Ac
11	Mr.	Donald	Ac

ภาพที่ 2: WebSQL

2.4 งานวิจัยที่เกี่ยวข้อง

มีผู้นำเสนอสถาปัตยกรรมแบบออฟไลน์ [6] โดยเพิ่มผู้ให้บริการแบบ Local เพื่อให้ผู้ใช้สามารถร้องขอแทนที่จะร้องขอตรงไปยังผู้ให้บริการจริง รวมถึงการนำเสนอโมดูลเคิร์ฟที่ไว้ติดตั้งบนโทรศัพท์เคลื่อนที่ [7] เพื่อให้ผู้ใช้สามารถเปิดเว็บได้โดยไม่ต้องเชื่อมต่อกับเครือข่าย โดยที่โมดูลเคิร์ฟนั้นจะทำหน้าที่เก็บข้อมูลเว็บไว้บนเครื่องของผู้ใช้ ซึ่งถ้าหากพิจารณาจะพบว่า ทั้ง 2 งานนั้นมีสิ่งที่เหมือนกันในการทำให้ออффไลน์ การทำงานในสภาวะออฟไลน์ก็คือการบันทึกข้อมูลส่วนหนึ่งไว้บนฝั่งผู้ใช้นั่นเอง

นอกจากนั้นมีการนำคุณลักษณะใหม่ของเอชทีเอ็มแอล 5 มาใช้ในการทำงานแบบออฟไลน์เช่น การประยุกต์ใช้ Web Storage ในการเก็บข้อมูลโจทย์คำถามของนักเรียนขณะอยู่ในโรงเรียน โดยที่นักเรียนสามารถนำกลับไปตอบคำถามที่บ้าน [8] และมีการนำเสนอสถาปัตยกรรมใหม่ที่ออกแบบให้เว็บประยุกต์สามารถทำงานได้บนเครือข่ายที่ไม่เสถียร [9]

3. วิธีดำเนินการวิจัย

ในงานวิจัยนี้ทำการศึกษาถึงวิธีการที่จะทำให้เว็บสามารถทำงานได้ในสภาวะออฟไลน์ ประกอบด้วย 2 ปัจจัยหลักคือการแคชไฟล์เว็บ และการเก็บข้อมูลบนฝั่งผู้ใช้

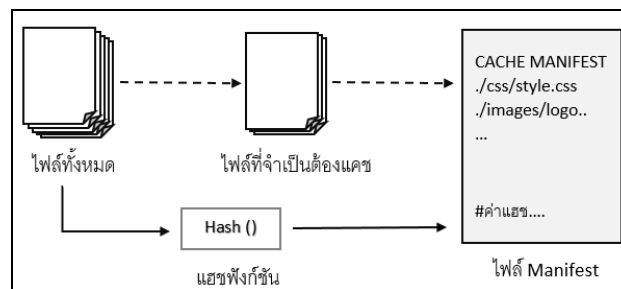
3.1 การแคชไฟล์และการควบคุมรุ่นของเว็บ

เว็บที่สามารถทำงานในสภาวะออฟไลน์ได้นั้นสิ่งแรกที่ต้องพิจารณาคือต้องสามารถเข้าใช้งานได้ตลอดเวลาที่ผู้ใช้ต้องการซึ่งโดยปกติหากอยู่ในสภาวะออฟไลน์

เริ่มจากการแบ่งประเภทของไฟล์เว็บออกเป็น 2 กลุ่มประกอบด้วย ไฟล์ที่จำเป็นต้องแคช และไฟล์ที่ไม่จำเป็นต้องแคช เพื่อรักษาความปลอดภัย และลดปริมาณการแคชที่ไม่จำเป็น ตัวอย่างไฟล์ที่จำเป็นต้องแคช เช่น รูปภาพ, CSS, JavaScript และไฟล์ที่ใช้ในการแสดงผล ส่วนไฟล์ที่ไม่จำเป็นต้องแคชไฟล์ที่ใช้เก็บการตั้งค่า ไฟล์ที่ใช้เก็บฟังก์ชันหรือคำสั่งการประมวลผลโปรแกรมต่าง ๆ จากนั้นนำรายการไฟล์ที่จำเป็นต้องแคชใส่ลงในไฟล์ Manifest และเนื่องจากการใช้งาน ApplicationCache เว็บเบราว์เซอร์จะไม่ทำการดึงข้อมูลไฟล์ที่ทำการแคชไปแล้วซ้ำเวลาที่ผู้ใช้เปิดเว็บใหม่ นอกจากจะมี การแก้ไขไฟล์ Manifest ซึ่งเมื่อเว็บเบราว์เซอร์พบว่าการแก้ไขไฟล์ Manifest ก็จะทำการดึงแคชเก่าทั้งหมด และทำการดึงแคชใหม่อีกครั้ง ส่งผลให้หากมีการปรับปรุงแก้ไขเว็บบนฝั่ง ผู้ให้บริการจะต้องทำการแก้ไขไฟล์ Manifest ด้วย

เมื่อมีการปรับปรุงแก้ไขเว็บส่งผลให้มีการเปลี่ยนแปลงรายการไฟล์ที่จำเป็นต้องแคช เช่น มีไฟล์รูปไฟล์ใหม่ หรือมีการลบไฟล์ออก ก็จะส่งผลให้เกิดการเปลี่ยนแปลงไฟล์ Manifest เมื่อผู้ใช้เข้าใช้งานเว็บเบราว์เซอร์ก็จะตรวจพบว่าไฟล์ Manifest มีการเปลี่ยนแปลงก็จะทำการดาวน์โหลดแคชและเก็บแคชใหม่ตามรายการไฟล์ที่กำหนดใน Manifest แต่ถ้าหากเป็นการปรับปรุงแก้ไขเว็บที่ไม่ส่งผลต่อรายการไฟล์ที่

จำเป็นต้องแคช เช่น การแก้ไขโปรแกรม หรือการเปลี่ยนแปลงรายการไฟล์ที่ไม่จำเป็นต้องแคช ไฟล์ Manifest ก็จะไม่มีการแก้ไข เมื่อผู้ใช้เข้าใช้งานเว็บเบราว์เซอร์ก็จะตรวจไม่พบ การเปลี่ยนแปลง และจะเข้าใช้ว่าใช้งานเว็บในรุ่นปัจจุบันอยู่แล้ว ดังนั้นจึงกำหนดวิธีในการควบคุมรุ่นของเว็บด้วยการนำแฮชฟังก์ชัน [10] มาช่วยตรวจสอบ โดยมีการทำงานดังภาพที่ 3



ภาพที่ 3: กระบวนการสร้างไฟล์ Manifest

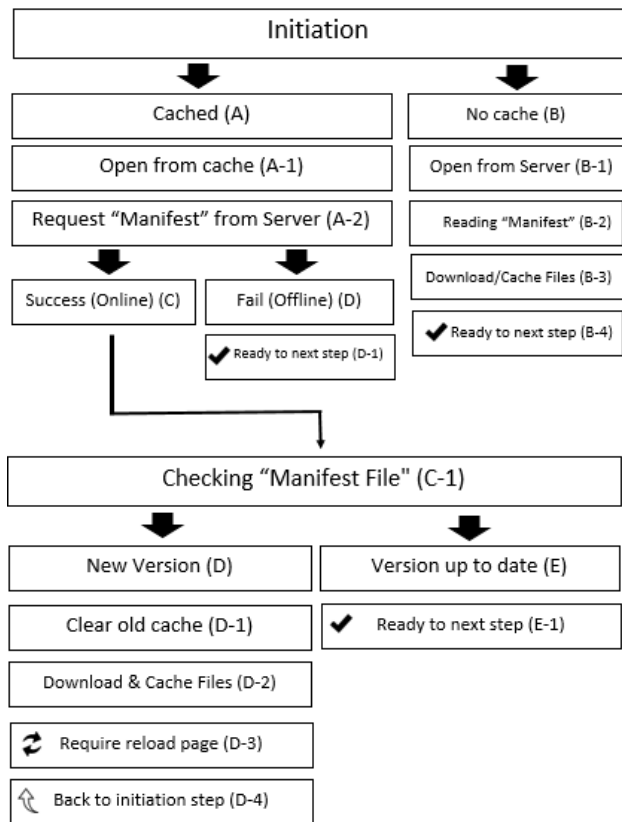
กระบวนการสร้างไฟล์ Manifest นั้นเริ่มจากการนำไฟล์ที่จำเป็นต้องแคชไปเขียนเป็นรายการไว้ใน CACHE MANIFEST จากนั้นนำเนื้อไฟล์ทั้งหมดเข้าแฮชฟังก์ชันทีละไฟล์ เมื่อได้ค่าแฮชนำไปเก็บไว้ แล้วนำไฟล์ใหม่เข้าแฮชฟังก์ชันต่อกับค่าแฮชก่อนหน้า โดยจะทำการวนการนี้ไปเรื่อย ๆ จนกว่าจะครบทุกไฟล์ ค่าแฮชที่ได้สุดท้ายจะนำไปใส่ไว้ในไฟล์ Manifest หลังเครื่องหมาย “#” เมื่อมีการแก้ไขเปลี่ยนแปลงเว็บ ค่าแฮชก็จะเปลี่ยนไปทำให้เว็บเบราว์เซอร์ในฝั่งผู้ใช้สามารถรู้ได้ว่าต้องทำการล้างแคชและเก็บแคชใหม่ตามรายการไฟล์ที่กำหนด

3.2 การเริ่มต้นและเตรียมความพร้อมเว็บ

หลังจากกระบวนการตรวจสอบความพร้อมใช้งานเอชทีเอ็มแอล 5, ApplicationCache และ WebSQL ของเว็บเบราว์เซอร์บนฝั่งผู้ใช้เป็นที่เรียบร้อย ก็จะเป็นกระบวนการเริ่มต้นและเตรียมความพร้อมเว็บเพื่อรองรับการทำงานในสภาวะออฟไลน์โดยจะมีกระบวนการดังภาพที่ 4

การเริ่มต้นและเตรียมความพร้อมเมื่อผู้ใช้เปิดเว็บจะทำการตรวจสอบว่ามีการแคชหรือไม่ (A, B) ในกรณีที่มีการแคชเรียบร้อยแล้วเว็บเบราว์เซอร์จะนำไฟล์ที่แคชไว้มาแสดงผล (A-1) โดยไม่มีการร้องขอไปยังผู้ให้บริการ จากนั้นจึงทำการร้องขอไฟล์ Manifest จากผู้ให้บริการ (A-2) หากการร้องขอไม่สำเร็จจะถือว่าอยู่ในสภาวะออฟไลน์ (D) และพร้อมที่จะทำงานในขั้นต่อไป (D-1) แต่หากการร้องขอสำเร็จ (C) จะทำ

การตรวจสอบไฟล์ Manifest (C-1) ว่ามีการเปลี่ยนแปลงหรือไม่



ภาพที่ 4: กระบวนการเริ่มต้นและเตรียมความพร้อมเว็บ

ถ้าไม่มีการเปลี่ยนแปลงจะถือว่าเว็บอยู่ในรุ่นปัจจุบัน (E) และพร้อมที่จะทำงานในขั้นต่อไป (E-1) แต่ถ้าหากมีการเปลี่ยนแปลงจะถือว่ามีการปรับรุ่นเว็บ (D) เว็บเบราว์เซอร์จะทำการล้างแคชเก่า (D-1) ดาวน์โหลดไฟล์มาแคชใหม่ (D-2) ตามรายการไฟล์ใน Manifest เมื่อการแคชเสร็จสมบูรณ์จะแจ้งให้ผู้ใช้ทำการเปิดเว็บใหม่อีกครั้ง (D-3) เมื่อทำการเปิดเว็บใหม่ก็จะเริ่มต้นกระบวนการเดิมอีกครั้ง (D-4)

ส่วนในกรณีที่ไม่มีแคชตั้งแต่ต้น (B) เว็บเบราว์เซอร์ก็จะทำการเปิดเว็บด้วยการร้องขอไปยังผู้ให้บริการ (B-1) จากนั้นจะทำการอ่านไฟล์ Manifest (B-2) แล้วทำการดาวน์โหลดไฟล์มาแคช (B-3) ตามรายการไฟล์ใน Manifest เมื่อการแคชเสร็จสมบูรณ์ก็พร้อมที่จะทำงานในขั้นต่อไป (B-4)

กระบวนการเริ่มต้นและเตรียมความพร้อมของเว็บนั้นจะมีเพียงขั้นตอน B-4, D-1 และ E-1 เท่านั้นที่จะยอมให้เริ่มต้นทำงานในขั้นต่อไป โดยในขั้นต่อไปจะเป็น การ

บริหารจัดการฐานข้อมูล WebSQL ภายในเว็บเบราว์เซอร์ของผู้ใช้

3.3 การบริหารจัดการฐานข้อมูล

เมื่อกระบวนการเริ่มต้นและเตรียมความพร้อมของเว็บเรียบร้อย จะเข้าสู่กระบวนการจัดการฐานข้อมูล WebSQL ภายในเว็บเบราว์เซอร์ของผู้ใช้

ในการใช้ฐานข้อมูล WebSQL นั้นจะใช้วิธีการเขียนคำสั่งต่างๆ ในการจัดการกับฐานข้อมูลผ่านทาง JavaScript โดยเริ่มตั้งแต่การสร้างฐานข้อมูล การสร้างตารางที่ต้องการ นั้นหมายความว่าทุก ๆ ครั้งที่มีการใช้งาน จำเป็นที่จะต้องตรวจสอบว่าเว็บเบราว์เซอร์ของผู้ใช้มีฐานข้อมูลรุ่นปัจจุบันหรือไม่

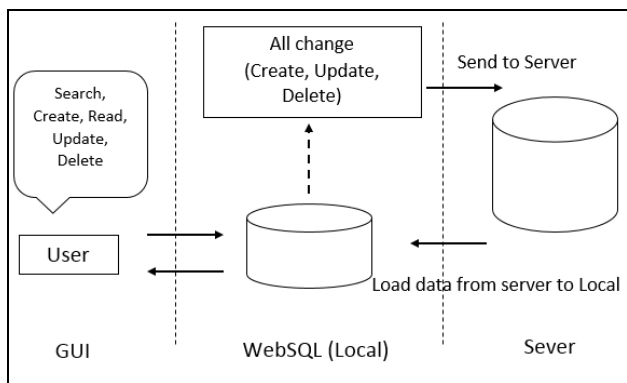
```

5
6   if(db.version == ""){
7       db.changeVersion("", "1", function(t){
8           t.executeSql("create table ...");
9       });
10  }
11
12  if(db.version == "1"){
13      db.changeVersion("1", "2", function(t){
14          t.executeSql("alter table ...");
15      });
16  }
    
```

ภาพที่ 5: การตรวจสอบรุ่นของฐานข้อมูล

และในการปรับรุ่นของฐานข้อมูลแต่ละครั้ง จะมีการสร้างตารางที่ต้องการอยู่อะไรบ้าง หรือมีการแก้ไขโครงสร้างตารางอะไรบ้าง ดังเช่นคำสั่งในภาพที่ 5

การจัดการกับข้อมูลจะเป็นกระบวนการที่จะควบคุม การเปลี่ยนแปลงข้อมูลของฝั่งผู้ใช้ และผู้ให้บริการให้มีข้อมูลที่ตรงกันและเป็นปัจจุบัน รวมถึงทำอะไรให้สามารถทำงานกับข้อมูลได้ในระหว่างที่อยู่ในสถานะออฟไลน์โดยมีการทำงานในภาพรวมดังนี้



ภาพที่ 6: การรับส่งข้อมูลระหว่างฝั่งผู้ใช้และผู้ให้บริการ

ผู้ใช้จะสามารถทำงานพื้นฐานทั้ง 5 กับข้อมูลประกอบด้วย การ ค้นหา, เพิ่ม, อ่าน, แก้ไข และลบได้ตลอดเวลาแม้ว่าจะอยู่ในสถานะออฟไลน์ เนื่องจากการทำงานนั้นจะกระทำกับข้อมูลบน WebSQL ที่อยู่ในฝั่งผู้ใช้เอง โดยที่จะมีกระบวนการเบื้องหลังที่จะคอยทำหน้าที่ส่งข้อมูลการเปลี่ยนแปลงที่เกิดขึ้นไม่ว่าจะเป็นการ เพิ่ม, แก้ไข และลบข้อมูลที่เกิดขึ้นใน WebSQL ไปยังผู้ให้บริการ และจะคอยดาวน์โหลดข้อมูลใหม่จากผู้ให้บริการมาเก็บไว้ใน WebSQL

ในกรณีที่มีการ เพิ่ม, แก้ไข หรือลบข้อมูล จะใช้วิธีการเดิมรูปแบบการแก้ไขไว้ที่ด้านหน้าของคีย์หลักของตารางข้อมูล เช่นหากมีการแก้ไขข้อมูล หรือลบข้อมูล ก็จะทำการเติม “Edit-” หรือ “Del-” ไว้ที่ด้านหน้าของคีย์หลัก ส่วนถ้าเป็นการเพิ่มใหม่ จะใช้คำ “New” เป็นคีย์หลัก ซึ่งคีย์หลักนี้จะสามารถมีค่าซ้ำได้ โดยตารางในฐานข้อมูลในฝั่งผู้ใช้จะเพิ่มคีย์หลักใหม่สำหรับใช้อ้างอิงข้อมูลภายในฝั่งผู้ใช้งานเอง โดยจะกำหนดชื่อฟิลด์ว่า “_id” โดยจะมีรูปการทำงานและการอ้างอิงการแก้ไขข้อมูลไปยังผู้ให้บริการดังตารางที่ 1 และตารางที่ 2

ตารางที่ 1: ตารางข้อมูลเมื่อมีการเปลี่ยนแปลงแก้ไขจากผู้ใช้

ID	Field1	Field2	_id
New	New Value	New Value	206
New	New Value	New Value	205
Edit-312	New Value	Value	204
Del-311	Value	Value	203
310	Value	Value	202
309	Value	Value	201

ตารางที่ 2: ตารางข้อมูลหลังจากส่งไปยังผู้ให้บริการและรับข้อมูลกลับมาบันทึกใหม่

ID	Field1	Field2	_id
314	New Value	New Value	205
313	New Value	New Value	204
312	New Value	Value	203
310	Value	Value	202
309	Value	Value	201

4. ผลการทดสอบ

ในการทดสอบจะใช้สถานการณ์จำลองทั้ง 5 ที่เป็นการทำงานพื้นฐานโดยทั่วไปประกอบด้วย การ ค้นหา, เพิ่ม, อ่าน, แก้ไข และลบข้อมูล หรือเรียกแบบย่อว่า SCRUD (Search, Create, Read, Update and Delete) [11], [12] โดยในการทดสอบนั้นจะทำในระหว่างที่อยู่ในสถานะออฟไลน์ ด้วยการถอดการเชื่อมต่อออกจากเครื่องที่ทดสอบ จากนั้นจึงเริ่มการทดสอบ ถ้าการทำงานพื้นฐานที่ทดสอบสำเร็จให้ทำการต่อการเชื่อมต่ออีกครั้งเพื่อให้มีสถานะออนไลน์ และทำการทดสอบว่าการเปลี่ยนแปลงของข้อมูลในระหว่างออฟไลน์ ได้ถูกส่งไปยังผู้ให้บริการอย่างถูกต้องหรือไม่ โดยผลการทดสอบพบว่าทุก ๆ การทำงานพื้นฐานสามารถทำงานได้เมื่ออยู่ในสถานะออฟไลน์ และเมื่อเชื่อมต่อกลับเข้าสู่สถานะออนไลน์ ข้อมูลที่เปลี่ยนแปลงระหว่างที่ออฟไลน์ ได้ถูกส่งไปยังผู้ให้บริการอย่างถูกต้อง

ตารางที่ 3: การทดสอบในสถานะออฟไลน์

การทำงาน	ขั้นตอนทดสอบ	ผลลัพธ์
ค้นหา	1. ผู้ใช้ใส่คำค้น 2. แสดงผลการค้นหา	สามารถทำได้
เพิ่ม	1. ผู้ใช้บันทึกข้อมูลใหม่ 2. ข้อมูลถูกบันทึก	สามารถทำได้ *
อ่าน	1. สามารถดูข้อมูล	สามารถทำได้
แก้ไข	1. ผู้ใช้แก้ไขข้อมูล 2. ข้อมูลถูกแก้ไข	สามารถทำได้ *
ลบ	1. ผู้ใช้ลบข้อมูล 2. ข้อมูลถูกลบ	สามารถทำได้ *

* ข้อมูลที่ถูกเพิ่ม, แก้ไข และลบจะถูกพักไว้บนฐานข้อมูลฝั่งผู้ใช้ และเมื่อการเชื่อมต่อกลับมาอยู่ในสถานะออนไลน์ก็จะส่งข้อมูลที่พักไว้ไปยังผู้ให้บริการ

5. สรุป

ผู้พัฒนาสามารถนำคุณลักษณะใหม่ของเอชทีเอ็มแอล 5 มาช่วยให้เว็บสามารถทำงานได้ในสถานะออฟไลน์ ด้วยการใช้อะปพลิเคชันแคชในการสนับสนุนการเข้าใช้เว็บในระหว่างที่อยู่ในสถานะออฟไลน์ และสามารถใช้อะปพลิเคชันแคชเป็นฐานข้อมูลที่สามารถเป็นที่พักข้อมูลชั่วคราวเพื่อรองรับการทำงานพื้นฐานทั่วไปได้ในระดับหนึ่งในสถานะออฟไลน์ อย่างไรก็ตามยังมีสิ่งที่ต้องพิจารณา หรือทำการศึกษาเพื่อหาวิธีในการจัดการที่ดีกว่า เช่น พิจารณาการเก็บข้อมูลลงในฐานข้อมูลฝั่งผู้ใช้ว่าสมควรเก็บข้อมูลอะไรบ้าง ต้องเก็บข้อมูลเท่าใด รวมถึงในแง่ความปลอดภัยของการเข้าถึงข้อมูลที่ถูกเก็บลงในฝั่งผู้ใช้ นอกจากนี้ข้อมูลฝั่งผู้ใช้ทำการเปลี่ยนแปลงในระหว่างที่ออฟไลน์เมื่อถูกส่งไปยังผู้ให้บริการหากเป็นข้อมูลที่ล้ำสมัย จะมีกระบวนการตรวจสอบอย่างไร ว่าเป็นข้อมูลที่สมควรแทนที่ลงในฐานข้อมูลจริงหรือไม่ เพราะถ้าบันทึกข้อมูลที่ล้ำสมัยแทนที่ลงไปก็อาจจะเป็นข้อมูลที่ไม่ถูกต้อง และอาจจะกระทบกับผู้ใช้อื่น

เอกสารอ้างอิง

- [1] Google, "Gears - Improving Your Web Browser," [Online] Availablefrom: <https://code.google.com/p/gears/> [2013,September 20].
- [2] Aaron Boodman, "Gears Team. Stopping the Gears," [Online] 2003. Availablefrom: <http://gearsblog.blogspot.com/2011/03/stopping-gears.html> [2013,September 20].
- [3] W3C, "Plan 2014," [Online] Availablefrom: <http://dev.w3.org/html5/decision-policy/html5-2014-plan.html> [2013,September 20].
- [4] W3C, "HTML5 Application Cache," [Online] Availablefrom: http://www.w3schools.com/html/html5_app_cache.asp [2013,September 20].
- [5] W3C, "Web SQL Database," [Online] Availablefrom: <http://www.w3.org/TR/webdatabase/> [2013,September 20].
- [6] Edgar E M Gonçalves and António Menezes Leitão, "Using common Lisp to prototype offline work in web applications for rich domains" *6th European Lisp Workshop*, pp. 18-27, 2009.
- [7] Yung-Wei Kao, et al, "An offline browsing system for mobile devices" *11th International Conference on Information Integration and Web-based Applications & Services*, pp. 338-343, 2009.
- [8] Royyana M Ijtihadie, et al, "Offline web application and quiz synchronization for e-learning activity for mobile browser" *TENCON 2010 - 2010 IEEE Region 10 Conference*, pp. 2402-2405, 2010.
- [9] CHEN Li-li and LIU Zheng-long, "Design of Rich Client Web Architecture Based on HTML5" *2012 Fourth International Conference on Computational and Information Sciences*, pp. 1009-1012, 2012.
- [10] Juric, M, "CUDA MD5 Hashing Experiments," [Online] Availablefrom: <http://majuric.org/software/cudamd5/> [2013,September 20].
- [11] Doug Serfass, et al, "Search, sort, pagination, association and (other) CRUD: a comparison of ActiveScaffold and MyFaces tomahawk" *SIGSOFT Software Engineering Notes*, 2008.
- [12] Bartalos Peter and Bieliková Mária, "(S)CRUD Pattern Support for Semantic Web Applications", *Slovak University of Technology in Bratislava*, 2008.

การพัฒนาต้นแบบเครื่องมือรวบรวมข่าวด้านการประกันวินาศภัย

A Development of a Non-life Insurance News Aggregator Prototype

จตุพิศ อัยราคอม (Jatupit Aiyarakom)¹ และ ทวิตตี้ เสนีวงศ์ ณ อยุธยา (Twittie Senivongse)²

ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

¹jatupit.a@student.chula.ac.th, ²twittie.s@chula.ac.th

บทคัดย่อ

ในแต่ละวันผู้บริหารในธุรกิจประกันวินาศภัยจำเป็นต้องรับข้อมูลข่าวสารจำนวนมากเพื่อให้ทราบสถานการณ์ความเป็นไปของธุรกิจ แต่ข้อมูลข่าวสารนั้นมีการเผยแพร่ในหลายช่องทาง ทั้งหนังสือพิมพ์ อีเมล และเว็บไซต์ อีกทั้งข่าวสารยังมีปริมาณมาก ทำให้การรวบรวมค้นหาข้อมูลเป็นไปได้ยาก บทความนี้นำเสนอต้นแบบเครื่องมือสำหรับรวบรวมข่าวสารด้านการประกันวินาศภัยจากเว็บไซต์ ได้แก่ ข้อมูลข่าวขององค์กรธุรกิจ ด้านประกันวินาศภัย องค์กรคู่ค้า องค์กรคู่แข่ง และองค์กรของรัฐที่เกี่ยวข้อง รวมถึงข่าวกฎหมายประกันวินาศภัย ต้นแบบของเครื่องมือสามารถรวบรวมเนื้อหาข่าวจากเว็บไซต์ที่ให้บริการข้อมูลเกี่ยวกับประกันวินาศภัย สกัดข้อมูลข่าว แล้วทำการจำแนกประเภทข่าวตามหมวดหมู่โดยอิงออนโทโลยีของโดเมนการประกันวินาศภัยซึ่งผู้พัฒนาได้สร้างขึ้น ผู้ใช้สามารถค้นคืนข้อมูลข่าวตามประเภทข่าวและตามคำหลักได้ จากการประเมินผล พบว่าต้นแบบเครื่องมือมีประสิทธิภาพค่อนข้างน่าพอใจ คือมีความแม่นยำในการจำแนกประเภทข่าว 0.81 และสามารถค้นคืนข่าวด้วยค่าความแม่นยำ 0.7 และค่าเรียกคืน 0.55

คำสำคัญ: เครื่องมือรวบรวมข่าว การจำแนก การค้นคืน

ออนโทโลยี การประกันวินาศภัย

Abstract

Every day, executives in non-life insurance business need to consume lots of information in order to know the current situations in the business. It is difficult to gather large amount of information that is distributed over different channels including newspapers, emails, and Web sites. This paper presents a prototype of a non-life insurance news aggregator that gathers information from

different Web sites. The gathered information includes news about a non-life insurance organization, its partners, its competitors, and a relevant governmental organization, as well as non-life insurance laws and regulations. The prototype can aggregate news contents from Web sites, and extract and classify news items based on the categories in the non-life insurance ontology that we propose. The prototype allows users to search news by categories and keywords. In an evaluation, the efficiency of the prototype is quite satisfactory, showing accuracy of 0.81 in news classification and precision of 0.7 and recall of 0.55 in news retrieval.

Keyword: News Aggregator, Classification, Retrieval, Ontology, Non-life Insurance

1. บทนำ

ในแต่ละวันผู้บริหารในธุรกิจประกันวินาศภัยจำเป็นต้องรับข้อมูลข่าวสารจำนวนมากเพื่อให้ทราบสถานการณ์ความเป็นไปของธุรกิจในปัจจุบัน ไม่ว่าจะเป็นข้อมูลบริษัทคู่แข่ง ข้อมูลคู่ค้า ข้อมูลประเภทการประกันภัยต่าง ๆ เป็นต้น จากกรณีศึกษาของบริษัทประกันวินาศภัยแห่งหนึ่งในประเทศไทย ในแต่ละวันการรวบรวมข้อมูลให้กับผู้บริหารทำโดยเจ้าหน้าที่ ซึ่งต้องรวบรวมจากหนังสือพิมพ์ อีเมล และแหล่งข้อมูลบนอินเทอร์เน็ต เช่น เว็บไซต์ และจัดกลุ่มข่าวให้กับผู้บริหาร การที่แหล่งข้อมูลที่เกี่ยวข้องมีอยู่หลายแหล่งและข้อมูลมีปริมาณมาก การรวบรวมข้อมูล จัดกลุ่ม และค้นหาข้อมูลที่ต้องการจึงเป็นไปได้ยาก จากปัญหาที่กล่าวมานี้ บทความนี้จึงนำเสนอการพัฒนาต้นแบบเครื่องมือรวบรวมข่าวด้านการประกันวินาศภัย โดยรวบรวมจากข้อมูลบนเว็บไซต์ที่เกี่ยวข้อง

ข้อมูลข่าวจะครอบคลุมข่าวขององค์กรธุรกิจด้านประกันวินาศภัย องค์กรคู่ค้า องค์กรคู่แข่ง และองค์กรของรัฐที่เกี่ยวข้อง รวมถึงข่าวกฎหมายประกันวินาศภัย จากนั้นนำข้อมูลข่าวมาวิเคราะห์เพื่อจำแนกหมวดหมู่ การจำแนกนั้นจะใช้วิธีจำแนกข้อความแบบจับคู่คำโดยตรง (Text Classification by Direct Match) [1] ซึ่งจำแนกโดยพิจารณาข้อความในข่าวเทียบกับคำศัพท์ในหมวดหมู่ตามออนโทโลยีของโดเมนการประกันวินาศภัยซึ่งผู้พัฒนาได้สร้างขึ้น ว่าสอดคล้องกับหมวดหมู่ใด ต้นแบบเครื่องมือจะสามารถค้นคืนข้อมูลข่าวประกันวินาศภัยตามประเภทข่าวที่จำแนกและตามคำหลักได้ จึงสามารถช่วยอำนวยความสะดวกในการรวบรวมและค้นคืนข่าวที่ผู้บริหารต้องการได้

ส่วนที่ 2 ของบทความจะกล่าวถึงงานวิจัยที่เกี่ยวข้อง ส่วนที่ 3 กล่าวถึงการพัฒนาต้นแบบเครื่องมือรวบรวมข่าวประกันวินาศภัยและตัวอย่างการใช้งาน ส่วนการประเมินผลการจำแนกข่าวและการค้นคืนจะอยู่ในส่วนที่ 4 และส่วนที่ 5 เป็นบทสรุป

2. งานวิจัยที่เกี่ยวข้อง

มีงานวิจัยหลายงานที่เกี่ยวข้องกับการจำแนกข้อมูล ร่วมกับการใช้ประโยชน์จากออนโทโลยีซึ่งหมายถึงกลุ่มของคำศัพท์และความสัมพันธ์ระหว่างคำศัพท์นั้นเพื่อใช้อธิบายโดเมนที่สนใจ [2] การจำแนกโดยทั่วไปนั้นทำโดยเทคนิคการจำแนกข้อความ (Text Classification) ซึ่งมีพื้นฐานอยู่บนการหาความคล้ายคลึงระหว่างข้อความกับกลุ่มคำศัพท์ที่บ่งบอกถึงหมวดหมู่ที่ต้องการจำแนก การวิเคราะห์ทำโดยคำนวณค่าน้ำหนัก Term Frequency-Inverse Document Frequency (TF-IDF) สำหรับคำต่าง ๆ ที่ปรากฏในข้อความและหมวดหมู่ เพื่อนำมาเปรียบเทียบค่าความคล้ายคลึงเชิงมุม (Cosine Similarity) [1]

งานวิจัยของ Goossen และคณะ [3] ทำการแนะนำข่าวให้ผู้ใช้โดยเสนอวิธีหาค่าน้ำหนักคำแบบ CF-IDF ซึ่งปรับปรุงจาก TF-IDF สำหรับการจำแนกข้อความข่าว โดยพิจารณาให้น้ำหนักเฉพาะแก่คำในข่าวที่ตรงกับคำศัพท์ในโปรไฟล์ของผู้ใช้ซึ่งระบุถึงความสนใจของผู้ใช้เท่านั้น แทนการคำนวณค่าน้ำหนักให้กับทุกคำในข่าวตามแบบ TF-IDF งานวิจัยของ Agarwal และคณะ [4] ทำการจำแนกข่าว RSS ตามประเภท

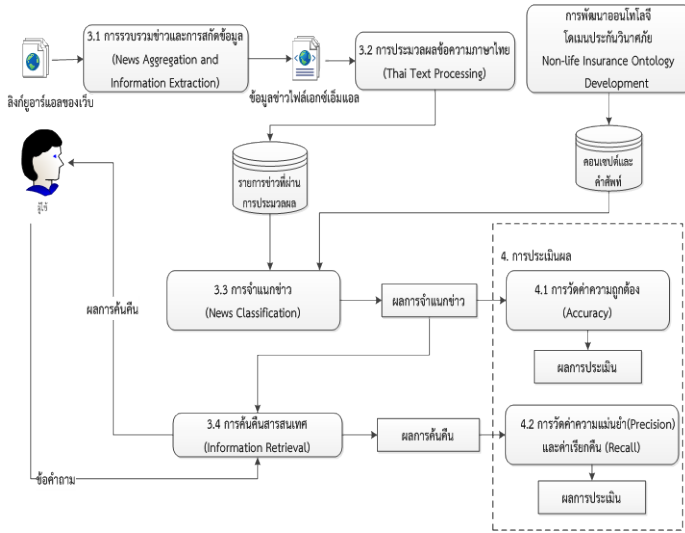
ข่าว โดยปรับวิธี CF-IDF ของ [3] ให้เป็น Weighted CF-IDF กล่าวคือมีการให้น้ำหนักเพิ่มแก่คำศัพท์ที่ปรากฏทั้งในหัวข้อข่าวและเนื้อข่าว เพราะถือว่ามีความสำคัญต่อการจำแนกมากกว่าคำที่พบเฉพาะในเนื้อข่าว ต่อมา Agarwal และคณะได้พัฒนาออนโทโลยีเฉพาะทางด้านการศึกษา [5] เพื่อใช้จำแนกและค้นคืนข่าวด้านการศึกษาในแบบเชิงความหมายได้ งานวิจัยของ Shi และ Liu [6] ใช้ประโยชน์จากออนโทโลยีในการประมวลผลข่าว โดยนำเสนอวิธีการในการสกัดข้อมูลจากเว็บไซต์ข่าวโดยการสร้างออนโทโลยีของโดเมนข่าวซึ่งเป็นไปตามโครงสร้างของหน้าเว็บข่าว จากนั้นจึงสร้างกฎในการสกัดข้อมูลตามออนโทโลยี เมื่อมีการเปลี่ยนรูปแบบในหน้าเว็บ จะใช้การปรับเปลี่ยนกฎในการสกัดไฟล์แทนการเปลี่ยนแปลงโปรแกรมที่ใช้ในการสกัดได้ แต่วิธีนี้ยังมีข้อจำกัดเพราะข้อมูลที่สกัดยังคงต้องเป็นไปตามส่วนประกอบตามโครงสร้างของหน้าเว็บตามออนโทโลยีอยู่ จึงไม่เหมาะกับต้นแบบเครื่องมือของผู้พัฒนาซึ่งจะใช้หน้าเว็บจากหลายแหล่งหลายโครงสร้าง นอกจากนี้ยังมีงานวิจัยที่ใกล้เคียงอื่น ๆ เกี่ยวกับการจำแนกข้อมูลหน้าเว็บและการค้นคืน เช่น งานวิจัยของ Roongkaew และ Prompoon [7] ได้ทำการพัฒนาเครื่องมือจำแนกข้อมูลหน้าเว็บของเครื่องมือทางวิศวกรรมซอฟต์แวร์ ว่าจัดเป็นเครื่องมือประเภทใดตามการแบ่งหมวดหมู่องค์ความรู้ด้านวิศวกรรมซอฟต์แวร์ พร้อมทั้งสามารถค้นคืนข้อมูลเครื่องมือได้

ผู้พัฒนาเห็นว่างานวิจัยข้างต้นสามารถนำมาใช้เป็นแนวทางในการพัฒนาต้นแบบเครื่องมือรวบรวมข่าวด้านการประกันวินาศภัยได้ โดยจะเน้นที่การพัฒนาออนโทโลยีเฉพาะทางสำหรับโดเมนการประกันวินาศภัยและใช้วิธีการสกัดข้อความข่าวในหลายรูปแบบจากหลายแหล่งเพื่อมาจำแนกและค้นคืน โดยการจำแนกจะใช้วิธีจับคู่คำโดยตรงซึ่งมีลักษณะเดียวกับวิธี CF-IDF ของ [3] นั่นเอง

3. ต้นแบบเครื่องมือรวบรวมข่าวด้านการประกันวินาศภัย

ภาพรวมของการพัฒนาต้นแบบเครื่องมือรวบรวมข่าวด้านการประกันวินาศภัยเป็นดังภาพที่ 1 ประกอบด้วย 5 กระบวนการหลักคือ 1) การรวบรวมข่าวและการสกัดข้อมูล (News Aggregation and Information Extraction) 2) การ

ประมวลผลข้อความภาษาไทย (Thai Text Processing) 3) การจำแนกข่าว (News Classification) ตามออนโทโลยีด้านการประกันวินาศภัย และ 4) การค้นคืนสารสนเทศ (Information Retrieval) 5) การประเมินผล รายละเอียดของแต่ละส่วนมีดังนี้



ภาพที่ 1: ภาพรวมของการพัฒนาต้นแบบเครื่องมือ

3.1 การรวบรวมข่าวและการสกัดข้อมูล

การรวบรวมเนื้อหาข่าวที่เกี่ยวข้องกับข้อมูลประกันวินาศภัยจะพิจารณาจากกรณีศึกษาของบริษัทประกันวินาศภัยแห่งหนึ่งในประเทศไทย โดยรวบรวมข้อมูลจาก 6 เว็บไซต์ ได้แก่

- 1) เว็บไซต์ข่าวเกี่ยวกับประกันวินาศภัย
<http://www.todayinsure.com/index.php>
- 2) เว็บไซต์สำนักงานคณะกรรมการกำกับและส่งเสริมการประกอบธุรกิจประกันภัย (คปภ.) สำหรับข้อมูลกฎหมายเกี่ยวกับประกันวินาศภัย
<http://www.oic.or.th/th/home/index.php>
- 3) เว็บไซต์บริษัทคู่แข่ง 2 ราย คือ
บริษัท วิริยะประกันภัย จำกัด
<http://www.bangkokinsurance.com/index.htm>
บริษัท กรุงเทพประกันภัย จำกัด (มหาชน)
<http://www.bangkokinsurance.com>
- 4) เว็บไซต์บริษัทคู่ค้า 2 ราย คือ
บริษัท มิตซูบิชิ มอเตอร์ส ประเทศไทย
<http://www.mitsubishi-motors.co.th/>
บริษัท ตรีเพชรอิซูซุเซลส์ จำกัด
<http://www.isuzu-tis.com/>

การรวบรวมทำโดยใช้เทคนิคการครอว์ลิงก์ (Crawling) ข้อมูลนำเข้าสำหรับการครอว์ลิงก์คือ ลิงก์ยูอาร์แอลของเว็บที่เกี่ยวข้องกับประกันวินาศภัย เมื่อทำการรวบรวมเนื้อหาข่าวจากหน้าเว็บต่าง ๆ แล้ว ผลลัพธ์ที่ได้คือไฟล์เอกซ์เอ็มแอลของเนื้อหาในหน้าเว็บ จากนั้นนำไฟล์เอกซ์เอ็มแอลมาสกัดเนื้อหาข่าวตามโครงสร้างของข่าว ได้แก่ หัวข้อข่าว คำอธิบายข่าว วันที่ประกาศข่าว แล้วทำการแปลงไฟล์เอกซ์เอ็มแอลเป็นไฟล์เอกซ์เอ็มแอล

3.2 การประมวลผลข้อความภาษาไทย

ประกอบด้วย 3 ขั้นตอนดังนี้

3.2.1 การตัดคำจากข้อความข่าว (News Tokenization)

เมื่อได้ข้อมูลข่าวตั้งต้นมาแล้ว จะทำการตัดคำโดยใช้คลาสไลบรารี (Class Library) ที่สนับสนุนโดยไอบีเอ็ม ชื่อว่า ICU4NET [8]

3.2.2 การจัดกลุ่มคำนาม (Noun grouping)

เป็นขั้นตอนตรวจสอบว่าคำใดที่เป็นคำเฉพาะทางของโดเมนประกันวินาศภัย และมีความสำคัญต่อการจำแนกข่าว เมื่อพบว่าการตัดคำไม่ถูกต้อง จะทำการจัดกลุ่มคำให้ถูกต้อง

3.2.3 การตัดคำหยุด (Stop-Word List Removal)

เป็นขั้นตอนการนำคำที่ไม่มีความสำคัญ ซึ่งเป็นคุณลักษณะที่ไม่เกี่ยวข้องหรือไม่มีประโยชน์ในการจำแนกข่าวออกไป ตัวอย่างคำหยุด เช่น ที่ ใน ว่า และ จะ มี ได้ ของ เป็นต้น ผลลัพธ์ที่ได้จากกระบวนการนี้คือ รายการข่าวที่ตัดคำที่ไม่มีนัยสำคัญออกแล้ว

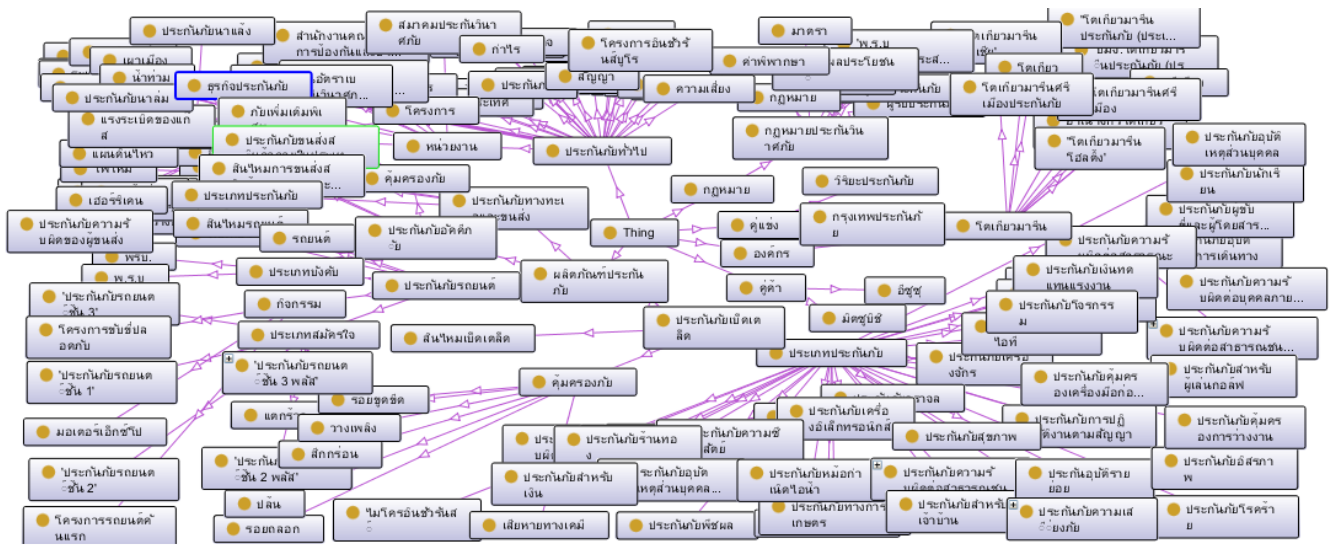
3.3 การจำแนกข่าว

ประกอบด้วย 2 ขั้นตอนดังนี้

3.3.1 การสร้างออนโทโลยีสำหรับโดเมนประกันวินาศภัย

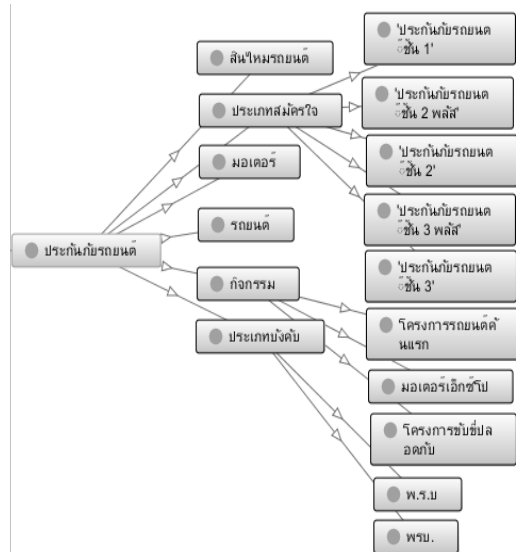
การสร้างออนโทโลยีจะใช้โปรแกรม Protégé [9] เพื่อกำหนดองค์ความรู้ของโดเมน โดยการกำหนดคอนเซปต์หรือคำศัพท์ที่เกี่ยวกับการประกันวินาศภัย โดยใช้แนวทางการกำหนดคอนเซปต์แบบผสมผสาน (Combination Development) กล่าวคือ นำความต้องการขององค์กรกรณีศึกษามานิยามคอนเซปต์ขององค์ความรู้ รวมทั้งรวบรวมคำศัพท์เกี่ยวกับประกัน

วิทยาศาสตร์และคำสำคัญที่ปรากฏในข่าว มานิชยามคำศัพท์ที่สัมพันธ์กัน ภาพรวมของออนโทโลยีที่ได้เป็นดังภาพที่ 2



ภาพที่ 2: ออนโทโลยีสำหรับโดเมนประกันวินาศภัย

ออนโทโลยีประกอบด้วยคำศัพท์หรือคอนเซปต์ใน 9 หมวดหมู่ ได้แก่ 1) ประกันภัยรถยนต์ 2) ประกันภัยทางทะเลและขนส่ง 3) ประกันภัยอสังหาริมทรัพย์ 4) ประกันภัยเบ็ดเตล็ด 5) ประกันภัยทั่วไป 6) องค์กร 7) คู่แข่ง 8) คู่ค้า และ 9) กฎหมายประกันวินาศภัย โดย 4 หมวดแรกนั้นเป็นการกำหนดหมวดหมู่ตามคปก. ผู้พัฒนาได้เพิ่มเติมคอนเซปต์ลงในหมวดหมู่ดังกล่าวพร้อมกับเสริมด้วยคอนเซปต์ในอีก 5 หมวดหมู่ รวมทั้งสิ้นมี 536 คอนเซปต์ ภาพที่ 3 แสดงตัวอย่างความสัมพันธ์ของคอนเซปต์ในหมวดประกันภัยรถยนต์ ซึ่งประกอบด้วยคอนเซปต์ลูกของ *ประกันภัยรถยนต์* ในสองลำดับชั้น และคอนเซปต์ *พ.ร.บ* เท่าเทียมกับคอนเซปต์ *พรบ.*



ภาพที่ 3: ตัวอย่างคอนเซปต์ในหมวดหมู่ประกันภัยรถยนต์

3.3.2 การจำแนกข่าว

การจำแนกข่าวจะใช้วิธีจำแนกข้อความแบบจับคู่คำโดยตรง (Text Classification by Direct Match) [1] ดังนี้

กำหนดให้ออนโทโลยีสำหรับโดเมนประกันวินาศภัย C คือเซตของ t คอนเซปต์:

$$C = \{c_1, c_2, \dots, c_t\} \text{ where } t = 536 \quad (1)$$

แต่ละหมวดหมู่ของออนโทโลยีการประกันวินาศภัย d_j คือเซตของ q คอนเซปต์ และข้อมูลข่าวจากหน้าเว็บ p หนึ่งชิ้นคือเซตของ s คอนเซปต์:

$$d_j = \{c_1^{d_j}, c_2^{d_j}, \dots, c_q^{d_j}\}, \text{ where } c^{d_j} \in C \text{ and } j = 1, \dots, 9 \quad (2)$$

$$p = \{c_1^p, c_2^p, \dots, c_s^p\}, \text{ where } c^p \in C \quad (3)$$

เมื่อพิจารณาแต่ละหมวดหมู่ของออนโทโลยีและข้อมูลข่าวจากหน้าเว็บเป็นเอกสารที่ต้องเปรียบเทียบความคล้ายคลึงกันเพื่อการจำแนก เราจะใช้ทฤษฎี Vector Space Model ในการสร้างเวกเตอร์ที่เป็นตัวแทนเอกสาร

กำหนดให้เวกเตอร์ \vec{d}_j เป็นตัวแทนของเอกสารหมวดหมู่ตามออนโทโลยี d_j :

$$\vec{d}_j = (w_{1,j}, w_{2,j}, \dots, w_{t,j}) \quad (4)$$

โดยที่ w_{ij} คือค่าน้ำหนักความสำคัญของคอนเซปต์ c_i ใน d_j
 โดยที่ $c_i \in C$ และ $w_{ij} \geq 0$ ค่า w_{ij} คำนวณโดย

$$w_{i,j} = \begin{cases} tf_{i,j} \times idf_i = (1 + \log_2 fi_{i,j}) \times \log_2 \frac{N}{n_i} & \text{if } fi_{i,j} > 0 \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

โดยที่ tf_{ij} = ค่าความถี่ (Term Frequency) ของคอนเซปต์ c_i ในเอกสาร d_j

idf_i = ค่าความถี่ในเอกสารแบบผกผัน (Inverse Document Frequency) ของคอนเซปต์ c_i

f_{ij} = จำนวนครั้งที่คอนเซปต์ c_i ปรากฏในเอกสาร d_j

f_{kj} = จำนวนครั้งที่คอนเซปต์ c_k ใด ๆ ปรากฏในเอกสาร d_j

N = จำนวนเอกสารทั้งหมด

n_i = จำนวนเอกสารที่คอนเซปต์ c_i ปรากฏ

จากนั้นกำหนดให้เวกเตอร์ \vec{p} เป็นตัวแทนของเอกสารข้อมูลข่าวจากหน้าเว็บ p :

$$\vec{p} = (w_{1,p}, w_{2,p}, \dots, w_{t,p}) \quad (6)$$

โดยที่ $w_{i,p}$ คือค่าน้ำหนักความสำคัญของคอนเซปต์ c_i ในเอกสารข่าวจากหน้าเว็บ p ซึ่งคำนวณแบบเดียวกับ w_{ij} โดยใช้สมการ (5)

เมื่อได้เวกเตอร์ทั้งหมดแล้ว การจำแนกข่าวจากหน้าเว็บว่าเกี่ยวข้องกับหมวดหมู่ใดของการประกันวินาศภัย จะทำโดยคำนวณค่าความคล้ายคลึงระหว่างแต่ละเอกสารหมวดหมู่ของออนไลน์กับเอกสารข่าวจากหน้าเว็บ โดยใช้การวัดค่าความคล้ายคลึงเชิงมุม (Cosine Similarity) [1]:

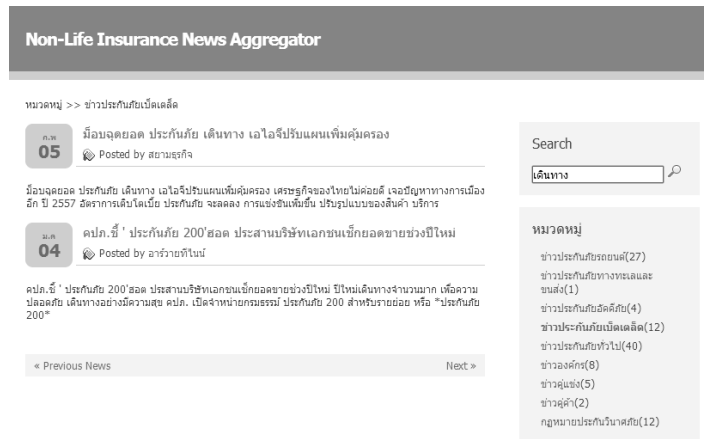
$$sim(d_j, p) = \frac{\vec{d}_j \cdot \vec{p}}{|\vec{d}_j| \times |\vec{p}|} = \frac{\sum_{i=1}^t w_{i,j} \times w_{i,p}}{\sqrt{\sum_{i=1}^t w_{i,j}^2} \times \sqrt{\sum_{i=1}^t w_{i,p}^2}} \quad (7)$$

ค่า $sim(d_j, p)$ อยู่ระหว่าง 0 ถึง 1 ข้อมูลข่าวจากหน้าเว็บจะถูกจำแนกอยู่ในหมวดหมู่ d_j ที่มีค่าความคล้ายคลึงนี้มากที่สุด

3.4 การค้นคืนสารสนเทศ

ต้นแบบเครื่องมือรวบรวมข่าวด้านการประกันวินาศภัยสนับสนุนการค้นคืนข่าวที่ได้จำแนกหมวดหมู่ไว้แล้ว ในรูปของเว็บแอปพลิเคชัน ผู้ใช้สามารถค้นคืนข่าวในหมวดที่สนใจหรือค้นคืนจากคำหลักซึ่งจะค้นจากทั้งในหมวดหมู่ข่าว หัวข้อข่าว และเนื้อหาข่าว การค้นคืนจะทำโดยใช้ประโยชน์จากออนไลน์ของโดเมนประกันวินาศภัยที่สร้างขึ้นมา ตัวอย่างเช่น

ในภาพที่ 4 เมื่อค้นคืนด้วยคำหลักว่า *เดินทาง* ต้นแบบเครื่องมือจะค้นหาคำหลักจากออนไลน์โทโลจี ซึ่งจะพบคอนเซปต์ *ประกันภัยเดินทาง* แต่เนื่องจากคอนเซปต์นี้เป็นคอนเซปต์ลูกของคอนเซปต์ *ประกันภัยเบ็ดเตล็ด* ข่าวที่เกี่ยวกับคำหลัก *เดินทาง* จึงจะถูกจำแนกในกระบวนการจำแนกให้เป็นข่าวในหมวดหมู่ *ประกันภัยเบ็ดเตล็ด* หรือหากผู้ใช้สนใจข่าวในหมวดหมู่ใดโดยเฉพาะก็สามารถเลือกหมวดหมู่จากด้านขวาของจอภาพ เช่น ดังภาพที่ 5 แสดงข่าวที่ได้จากการเลือกหมวดหมู่ *ประกันภัยรถยนต์*



ภาพที่ 4: ตัวอย่างผลการค้นคืนข้อมูลข่าว



ภาพที่ 5: ตัวอย่างผลการจำแนกข้อมูลข่าวตามหมวดหมู่

4. การประเมินผล

การประเมินต้นแบบเครื่องมือรวบรวมข่าว ประกอบด้วย การประเมินประสิทธิภาพการจำแนกข่าวจากหน้าเว็บ และการประเมินประสิทธิภาพการค้นคืน โดยใช้ข้อมูลข่าวจาก 6 เว็บไซต์ตามที่ระบุในหัวข้อที่ 3.1 ซึ่งรวบรวมในช่วงวันที่ 21-25 กุมภาพันธ์ 2557 จำนวนรวม 200 ข่าว

4.1 การประเมินผลการจำแนกข่าว

การประเมินผลการจำแนกข่าว 200 ข่าวทำโดยวัดค่าความถูกต้อง (Accuracy) เปรียบเทียบกับผลการจำแนกของผู้เชี่ยวชาญโดยใช้:

$$accuracy = \frac{\text{number of news items correctly classified by tool}}{\text{number of news items classified by tool}} \quad (8)$$

ต้นแบบเครื่องมือมีค่าความถูกต้องของการจำแนกเป็น 0.81 โดยข่าวที่จำแนกไม่ถูกต้องมีสาเหตุหลักจากการตัดคำภาษาไทย โดยเฉพาะคำเฉพาะและคำที่เขียนทับศัพท์ ตัวอย่างเช่น คำว่า *โตเกียวมาริน* จะตัดคำได้เป็น *โตเกียว มา* และ *ริน* ในขณะที่คำว่า *มา* จะเป็นคำหยุดจึงถูกตัดออก เป็นต้น ผู้พัฒนาได้ทำการพิจารณาจัดกลุ่มคำเหล่านี้ แต่ไม่สามารถครอบคลุมได้ทุกกรณี

4.2 การประเมินผลการค้นคืนข่าว

การประเมินผลการค้นคืนทำโดยค้นคืน 200 ข่าว ด้วยข้อความ (Query) 7 รายการ แล้ววัดค่าความแม่นยำ (Precision) และค่าเรียกคืน (Recall) เปรียบเทียบกับผลการค้นคืนของผู้เชี่ยวชาญโดยใช้:

$$precision = \frac{\text{number of relevant news items retrieved by tool}}{\text{number of news items retrieved by tool}} \quad (9)$$

$$recall = \frac{\text{number of relevant news items retrieved by tool}}{\text{number of relevant news items}} \quad (10)$$

ต้นแบบเครื่องมือสามารถค้นคืนด้วยค่าความแม่นยำเป็น 0.7 และค่าเรียกคืนเป็น 0.55 โดยข่าวที่ค้นคืนไม่ถูกต้องมีสาเหตุหลักจากการที่ต้นแบบเครื่องมือไม่ได้ทำสเต็มมิง (Stemming) เพื่อลดรูปของคำให้เป็นรากศัพท์

5. บทสรุป

ต้นแบบเครื่องมือรวบรวมข่าวด้านการประกันวินาศภัยสามารถรวบรวมข่าวสารด้านการประกันภัยวินาศภัยจากเว็บไซต์ที่องค์กรสนใจ และสามารถจำแนกและค้นคืนข่าวได้

โดยมีประสิทธิภาพอยู่ในระดับที่ค่อนข้างเป็นที่พอใจ จึงเป็นประโยชน์ต่อการรับรู้ความเคลื่อนไหวในวงการและสนับสนุนการวางแผนการดำเนินงานขององค์กรธุรกิจประกันวินาศภัยต่อไป อย่างไรก็ตามเครื่องมือนี้ยังถือว่าเป็นเพียงต้นแบบเนื่องจากยังสามารถปรับปรุงเพิ่มเติมได้ในหลายประเด็น เช่น เพิ่มการทำสเต็มมิง และใช้วิธีการจำแนกข้อความที่มีประสิทธิภาพมากขึ้น นอกจากนี้ออนโทโลยีด้านการประกันวินาศภัยที่เสนอยังมีความเจาะจงกับองค์กรกรณีศึกษาอยู่ โดยหากองค์กรอื่นต้องการนำไปใช้ ควรที่จะปรับแต่งออนโทโลยี โดยเฉพาะในส่วนที่เกี่ยวกับองค์กร คู่แข่ง คู่ค้า ควรต้องเพิ่มคอนเซปต์ที่เกี่ยวข้องให้เหมาะสม อีกทั้งในการใช้งานจริงในธุรกิจ ควรที่จะพัฒนาการรวบรวมข่าวให้เป็นไปอย่างอัตโนมัติ โดยทำการครอว์ลิงก์ทุกวันเพื่อรวบรวม สกัดข้อมูล และจำแนกได้อย่างอัตโนมัติ เพื่อให้มีข้อมูลสำหรับค้นคืนอย่างเป็นปัจจุบันเสมอ รวมถึงรวบรวมข้อมูลจากเว็บไซต์ที่เกี่ยวข้องกับการประกันวินาศภัยให้ครอบคลุมทั่วถึงยิ่งขึ้น

เอกสารอ้างอิง

- [1] R. Baeza-Yates and B. Ribeiro-Neto, *Modern Information Retrieval, Second edition*. Essex: Addison Wesley, 2011, pp. 290-291, 77-78.
- [2] T. Gruber, *Ontology*, <http://tomgruber.org/writing/ontology-definition-2007.htm>
- [3] F. Goossen, W. IJntema, F. Frasincar, F. Hogenboom, and U. Kaymak, "News personalization using the CF-IDF semantic recommender," *Proc. Int. Conf. Web Intelligence, Mining and Semantics (WIMS '11)*, 2011.
- [4] S. Agarwal, A. Singhal, and P. Bedi, "Classification of RSS feed news items using ontology," *Proc. 12th Int. Conf. Intelligent Systems Design and Applications (ISDA 2012)*, 27-29 Nov 2012, pp. 491-496.
- [5] S. Agarwal, A. Singhal, and P. Bedi, "IPTC based ontological representation of educational news RSS feeds," *Proc. 3rd Int. Conf. Trends in Information, Telecommunication and Computing, Lecture Notes in Electrical Engineering*, Volume 150, 2013, pp. 353-359.
- [6] J. Shi and L. Liu, "Web information extraction based on news domain ontology theory," *Proc. IEEE 2nd Symp. Web Society (SWS 2010)*, 16-17 Aug 2010, pp. 416, 419.

- [7] W. Roongkaew and N. Prompoon, "Software engineering tools classification based on SWEBOK taxonomy and software profile," *Proc. 2nd Int. Conf. Informatics and Applications (ICIA 2013)*, 23-25 Sep 2013, pp. 122-128.
- [8] Google Code, *ICU4NET*,
<https://code.google.com/p/icu4net/>
- [9] Stanford Center for Biomedical Informatics Research,
Protégé, <http://protege.stanford.edu/>

ระบบสร้างกรณีทดสอบอัตโนมัติด้วยต้นไม้จำแนกข้อมูล Automated Generation of Test Cases System by Using Classification Tree

สุชาดา รัตน์กนเณตร (Suchada Ratanakongnate)¹ และกิตติชัย โสภิตเวศย์มนตรี (Kittichai Sopitwetmontree)²

ภาควิชาวิทยาการคอมพิวเตอร์และสารสนเทศ คณะวิทยาศาสตร์ประยุกต์

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

suchada.rata@gmail.com¹, ks_hok@hotmail.com²

บทคัดย่อ

บทความนี้กล่าวถึงกรอบแนวคิดการพัฒนาซอฟต์แวร์เพื่อช่วยลดระยะเวลาในการทดสอบ อีกทั้งยังช่วยให้การทดสอบครอบคลุมมากขึ้น โดยใช้หลักการต้นไม้จำแนกข้อมูล (Classification Tree) และอนุกรมวิธาน (Taxonomy) กรณีทดสอบอัตโนมัติถูกสร้างจากยูสเคสเดสคริปชัน (Use Case Description) ในขั้นตอนการออกแบบระบบ ซึ่งสามารถช่วยให้นักวิเคราะห์ระบบ นักทดสอบระบบและลูกค้าสามารถเห็นผลการทดสอบที่จะเกิดขึ้นหลังการพัฒนาระบบเสร็จสิ้นตามความต้องการ ทำให้สามารถเข้าใจตรงกันถึงภาพรวมของสิ่งที่เกิดขึ้น และยังสามารถวิเคราะห์ความครอบคลุมความสอดคล้องของกรณีทดสอบกับระบบอีกด้วย การที่สามารถตรวจสอบความถูกต้องได้ในช่วงเวลาของการวิเคราะห์ระบบ ทำให้สามารถแก้ไขการออกแบบระบบหรือกรณีทดสอบเพื่อตอบสนองความต้องการกับระบบที่มีอยู่ได้อย่างถูกต้องและทันท่วงที และยังมีส่วนช่วยในการประหยัดเวลาและค่าใช้จ่ายในการทดสอบระบบในกระบวนการผลิตซอฟต์แวร์

คำสำคัญ: การทดสอบซอฟต์แวร์ การวิเคราะห์ระบบ ยูสเคสกรณีทดสอบ ต้นไม้จำแนกข้อมูล อนุกรมวิธาน

Abstract

This paper is to develop a system to help reduce the duration of the testing. It also allows for a more comprehensive test using the classification tree and taxonomy. Automated test cases are created from the Use Case Description in the process of system design, which can help the system analysts, testers and customers to see the result for testing after the development will be

completed. In order to be more understand an overview of system requirement and also analyze the coverage of consistency of the test case to the system as well. That can be checked for accuracy at the time of analysis. In order that it is possible to modify the design or test cases to meet the requirements for the existing system correctly. It also helps in saving the time and expense of testing in the software process.

Keyword: Software Testing, System Analysis, Use Case, Test case, Classification Tree, Taxonomy

1. บทนำ

การทดสอบซอฟต์แวร์เป็นขั้นตอนที่สำคัญในวงจรการพัฒนาซอฟต์แวร์ที่ช่วยให้ซอฟต์แวร์ที่ถูกพัฒนามีคุณภาพและน่าเชื่อถือ นักวิเคราะห์ นักพัฒนาและนักทดสอบระบบมีส่วนรับผิดชอบถึงคุณภาพซอฟต์แวร์ที่ส่งถึงลูกค้า ทุกฝ่ายในวงจรการพัฒนาซอฟต์แวร์ควรมีส่วนร่วมในการช่วยทำให้คุณภาพซอฟต์แวร์มีคุณภาพมากขึ้น ปัญหาที่เกิดขึ้นเป็นประจำคือกรณีทดสอบไม่สามารถออกแบบให้ครอบคลุมกับซอฟต์แวร์ที่ถูกพัฒนาออกมารวมทั้งยังใช้เวลานานเกินไป ทำให้ประสิทธิภาพการทำงานของคนในองค์กรลดลงด้วย

งานวิจัยนี้ได้นำเสนอวิธีในการแก้ปัญหาเรื่องความครอบคลุมและลดระยะเวลาการสร้างกรณีทดสอบโดยใช้วิธีการต้นไม้จำแนกข้อมูลและอนุกรมวิธาน ผสมกับการสร้างยูสเคสเดสคริปชัน ซึ่งเป็นขั้นตอนในการออกแบบระบบ เนื่องจากขั้นตอนการออกแบบยูสเคสเดสคริปชันเป็นขั้นตอนที่นักวิเคราะห์เริ่มมองเห็นภาพของระบบและลักษณะการทำงานของแต่ละฟังก์ชันชัดเจนพอสมควร อีกทั้งยังเป็นขั้นตอนที่นัก

ทดสอบสามารถร่วมพิจารณาโครงสร้างระบบซึ่งเข้าใจได้ไม่ยากเนื่องจากสามารถใช้แผนภาพยูสเคสและมีคำอธิบายการทำงานจากยูสเคสเดสคริปชัน งานวิจัยนี้จะทำการสร้างกรณีทดสอบจากยูสเคสที่ได้จากการออกแบบระบบ ยูสเคสเดสคริปชันคือหัวใจหลักของการสร้างกรณีทดสอบจากงานวิจัยนี้

ปัจจุบันมีวิธีการสร้างกรณีทดสอบอัตโนมัติจากหลายโมเดล งานวิจัยนี้ใช้หลักการต้นไม้จำแนกข้อมูลเพราะสามารถแบ่งแยกออกเป็นลักษณะของข้อมูล ทำให้สามารถวิเคราะห์กรณีทดสอบที่สร้างขึ้นในรูปแบบทรี เห็นภาพได้ชัดเจนขึ้น และสามารถควบคุมการสร้างกรณีทดสอบได้ตั้งแต่ความครอบคลุมที่น้อยที่สุดจนถึงมากที่สุด สามารถสร้างจากสายทางเดินที่เป็นไปได้จากยูสเคสเดสคริปชันซึ่งเป็นการอธิบายฟังก์ชันการทำงานในรูปแบบซินนารีโอ ทำการผสมผสานอนุกรมวิธานเพื่อจัดโครงสร้างข้อมูลในระบบให้เป็นสัดส่วนเพื่อความง่ายต่อการปรับปรุงแก้ไขระบบเพื่อใช้ในการทดสอบแบบถดถอย(Regression Testing) และค้นหาข้อมูลในระบบ อีกทั้งยังสร้างกฎฟังพาทอาศัยกัน เพื่อช่วยสร้างกรณีทดสอบที่ครอบคลุมเหมาะสมซึ่งช่วยลดทอนกรณีทดสอบที่ไม่จำเป็นต้องทดสอบทุกกรณีเพื่อเป็นการลดระยะเวลาในการทดสอบอีกด้วย

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

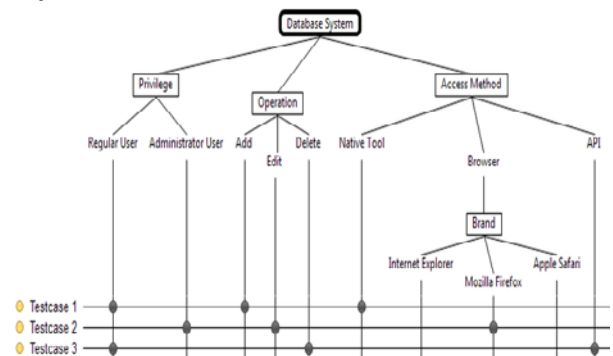
2.1 วิธีการต้นไม้จำแนกข้อมูล (Classification Tree Method)

วิธีการต้นไม้จำแนกข้อมูลคือหลักการที่ใช้ในการออกแบบกรณีทดสอบแบบ โดยอาศัยการจำแนกและจัดกลุ่มอินพุตที่เป็นไปได้ภายใต้ลักษณะการกระทำที่ผู้ใช้สามารถกระทำกับระบบได้ ถูกพัฒนาโดย Grimm และ Grochtmann ในปี 1933 วิธีการมี 2 ขั้นตอนดังนี้

1. ระบุพร้อมจัดประเภท โดเมนของอินพุตที่เกี่ยวข้องกับการทดสอบ (เรียกว่า Classifications) และระบุค่าที่เป็นไปได้ของโดเมนนั้น (เรียกว่า Classes)
2. รวบรวมโดเมนของอินพุตในระบบเพื่อสร้างกรณีทดสอบ

ตัวอย่างการสร้างต้นไม้จำแนกข้อมูลของระบบฐานข้อมูล ดังภาพที่ 1 ระบบฐานข้อมูลซึ่งแบ่งลักษณะของการกระทำออกเป็น 3 ส่วน ส่วนแรกคือ ลักษณะของผู้ใช้ในระบบ ส่วนที่สองคือการกระทำกับระบบ ส่วนที่สามคือ วิธีการเข้าถึงการใช้

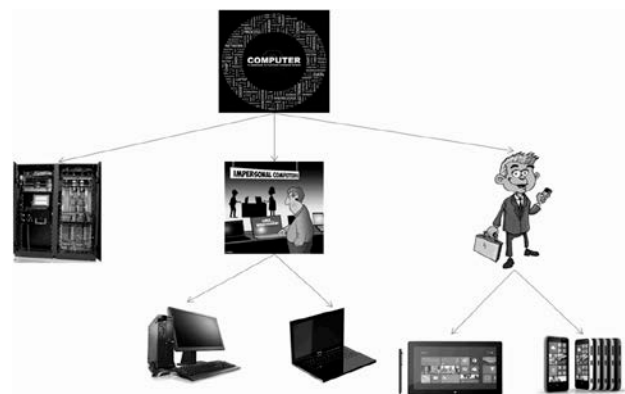
งานระบบ ยกตัวอย่างกรณีทดสอบที่ 1 คือผู้ใช้ทั่วไปทำการเพิ่มข้อมูลในระบบโดยการเข้าถึงโปรแกรมผ่านเครื่องมือทั่วไป



ภาพที่ 1: ตัวอย่างต้นไม้จำแนกข้อมูล

2.2 อนุกรมวิธาน (Taxonomy)

อนุกรมวิธานคือหลักการจัดจำแนกสิ่งต่างๆออกเป็นประเภท เพื่อแสดงความสัมพันธ์เป็นลำดับชั้น ซึ่งมีโครงสร้างเป็นรูปแบบต้นไม้ ซึ่งทำให้ง่ายต่อการจัดเก็บข้อมูลให้เป็นระเบียบ เพื่อง่ายต่อการดึงข้อมูลมาใช้หรือสะดวกต่อการค้นหา ทั้งยังง่ายต่อการออกแบบการเขียนโปรแกรมสนับสนุนได้มากขึ้น ดังภาพที่ 2 คือตัวอย่างอนุกรมวิธานของระบบคอมพิวเตอร์ ระบบคอมพิวเตอร์แบ่งได้เป็น เมนเฟรมคอมพิวเตอร์ คอมพิวเตอร์ส่วนบุคคล และคอมพิวเตอร์พกพา ในส่วนของคอมพิวเตอร์ส่วนบุคคลสามารถแยกสายวิวัฒนาการได้เป็น คอมพิวเตอร์ตั้งโต๊ะและคอมพิวเตอร์โน้ตบุ๊ก สุดท้ายคือคอมพิวเตอร์พกพามีสายวิวัฒนาการคือ แท็บเล็ตและสมาร์ตโฟน ซึ่งการใช้อนุกรมวิธานเพื่อที่จะนำกรณีทดสอบมาใช้ซ้ำ (Reuse Test Case) ในการทดสอบแบบถดถอย โดยไม่ต้องคิดและวิเคราะห์อีก



ภาพที่ 2: ตัวอย่างอนุกรมวิธาน

2.3 กฎที่พึ่งพาอาศัยกัน (Dependency Rule)

กฎการพึ่งพาอาศัยกันคือหลักการที่ใช้ในการสร้างความขึ้นต่อกันของอินพุตที่อยู่ภายใต้การกระทำที่แตกต่างกันหรือคลาสกับระบบ ซึ่งเป็นหลักการที่นิยมใช้ภายใต้หลักการต้นไม้จำแนกข้อมูล โดยมีวิธีการสร้างกฎหลากหลายวิธี ในงานวิจัยนี้ได้้นำวิธีการจาก [1] และ [2] มาประยุกต์ใช้ โดยถูกกำหนดไว้ 3 แบบดังนี้

1. ความสัมพันธ์แบบ OR คือความสัมพันธ์เริ่มต้นที่ตั้งค่าให้กับทุกอินพุตภายใต้ต้นละคลาสิกกัน ความสัมพันธ์นี้อนุญาตให้ทุกคลาสสามารถอยู่ในกรณีทดสอบเดียวกันได้ เช่น การถอนเงิน 100 บาท สามารถทำการตรวจสอบยอดเงินหรือไม่ตรวจสอบยอดเงินได้
2. ความสัมพันธ์แบบ NOT คือความสัมพันธ์ที่เกี่ยวข้องกัน โดยหากเกิดเหตุการณ์ที่ 1 แล้วเหตุการณ์ที่ 2 ไม่สามารถเกิดขึ้นได้ เช่น การกรอกรหัสผ่านไม่ถูกต้อง จะไม่สามารถทำการระบุจำนวนเงินในการถอนหรือโอนได้
3. ความสัมพันธ์แบบ AND คือความสัมพันธ์ที่เกี่ยวข้องกัน โดยหากเกิดเหตุการณ์ที่ 1 เหตุการณ์ที่ 2 ต้องเกิดขึ้น ลักษณะความสัมพันธ์นี้เกิดขึ้นไม่มากนักในงานวิจัยนี้ กำหนดไว้สำหรับกรณีทดสอบบางกรณี เช่น ธนาคารแห่งหนึ่งกำหนดกฎไว้ว่าหากผู้ใช้ทำการโอนเงินเข้าบัญชีใดก็ตามจะต้องทำการพิมพ์ใบเสร็จเพื่อเป็นหลักฐานทุกครั้ง ดังนั้นระบบห้ามละเว้นการพิมพ์ใบเสร็จไม่ว่ากรณีใดก็ตาม

2.4 กรอบแนวคิดการสร้างกรณีทดสอบจากคุณลักษณะความต้องการในรูปแบบฟอร์ม

งานวิจัยนี้ได้ดำเนินงานกรอบแนวคิดการสร้างกรณีทดสอบจากความต้องการในรูปแบบฟอร์ม [3] มาประยุกต์ใช้ในการพัฒนาซึ่งได้เสนอแนวคิดการสร้างกรณีทดสอบอัตโนมัติในขั้นตอนการสอบถามความต้องการ โดยสร้างแบบฟอร์มใช้ในการสอบถามความต้องการจากลูกค้า จากนั้นวิเคราะห์และสร้างกรณีทดสอบ ซึ่งเป็นการสร้างกรณีทดสอบจากความต้องการ

2.5 Test Sequence Generation from Classification Trees

ศึกษาและประยุกต์งานวิจัย [1], [2] โดยนำหลักการสร้างกฎ

ที่พึ่งพาอาศัยกันมาประยุกต์ใช้กับแนวทางในงานวิจัย อีกทั้ง [1], [2] ทำการสร้างกรณีทดสอบในขั้นตอนการออกแบบเช่นกันคือใช้สเตทชาร์ตไดอะแกรม เพื่อสร้างกรณีทดสอบตามลำดับเหตุการณ์ที่เกิดขึ้น

2.6 Test Case Design by Means of CTE XL

ศึกษาและประยุกต์งานวิจัย [8] ซึ่งได้ทำการทดลองและวิเคราะห์แล้วว่าจำนวนกรณีทดสอบที่ครอบคลุมเหมาะสมคือจำนวนกรณีทดสอบที่อยู่ระหว่างจำนวนกรณีทดสอบแบบครอบคลุมต่ำสุดและจำนวนกรณีทดสอบแบบครอบคลุมสูงสุด

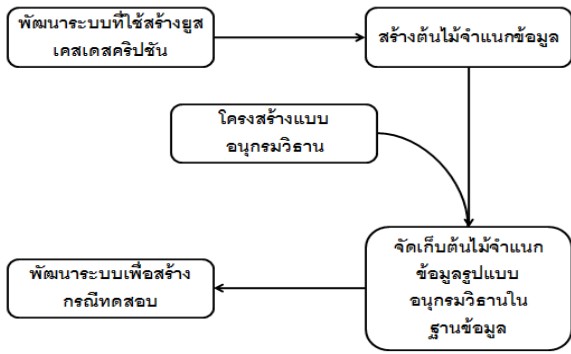
2.7 Automated Test Case Generation using Classification Trees

ศึกษางานวิจัย [11] ซึ่งทำการสร้างกรณีทดสอบโดยใช้ต้นไม้จำแนกข้อมูลที่มีการกำหนดค่าความน่าจะเป็นเพื่อใช้เป็นน้ำหนักในแต่ละประเภทโดเมนของอินพุต และค่าที่เป็นไปได้ของโดเมนนั้น ทำให้สามารถสร้างกรณีทดสอบตามลำดับความสำคัญของชุดกรณีทดสอบได้

3. วิธีดำเนินการวิจัย

3.1 ออกแบบขั้นตอนการทำงานวิจัย

ทำการออกแบบขั้นตอนการทำงานวิจัยดังภาพที่ 3 โดยเริ่มต้นสร้างระบบที่ใช้ในการสร้างยูสเคสเดสคริปชันซึ่งจะเกิดหลังขั้นตอนการสร้างแผนภาพยูสเคส โดยระบบนี้จะถูกพัฒนาขึ้นเพื่อสนับสนุนการกรอกข้อมูลในการอธิบายการทำงานของแต่ละฟังก์ชันการทำงานในยูสเคส ซึ่งมีส่วนประกอบสำคัญให้เห็นวิเคราะห์ทำการกรอกข้อมูลคือส่วนหลักในการทำงานของฟังก์ชันซึ่งคือการปฏิสัมพันธ์กันระหว่างผู้ใช้และระบบ ส่วนรองที่ขาดไม่ได้เช่นกันคือส่วนข้อยกเว้นที่เกิดขึ้นในฟังก์ชัน ซึ่งคือส่วนที่อธิบายการทำงานที่ไม่เป็นไปตามการทำงานของระบบ หลังจากนั้นพัฒนาระบบรองรับการจัดเก็บคลาสสิฟิเคชันต่างๆในรูปแบบอนุกรมวิธาน รวมถึงสนับสนุนการระบุคลาสหรืออินพุตข้อมูลที่เป็นไปได้ในแต่ละคลาสสิฟิเคชันของต้นไม้จำแนกข้อมูล และส่วนสุดท้ายคือพัฒนาระบบในการสร้างกรณีทดสอบอัตโนมัติ



ภาพที่ 3: โครงสร้างวิธีการทำงานของงานวิจัย

3.2 ใช้หลักอนุกรมวิธานออกแบบการจัดเก็บโดเมนอินพุต

ทำการออกแบบการจัดเก็บโดเมนอินพุตออกเป็นหมวดหมู่ในรูปของอนุกรมวิธาน โดยประยุกต์ใช้ในโครงสร้างฐานข้อมูล SQL เพื่อใช้ในการจัดเก็บ ค้นหาและพัฒนาระบบสร้างกรณีทดสอบดังภาพที่ 4

ตัวอย่างโดเมนอินพุต โดเมนเอทีเอ็ม แบ่งออกเป็น โดเมนการถอนเงิน โดเมนการโอนเงิน โดเมนการตรวจสอบยอดเงิน โดเมนของการกรอกรหัสผ่าน โดเมนการออกไปเสร็จ

ตัวอย่างโดเมนอินพุตระดับลูกของโดเมนการถอนเงินเอทีเอ็มคือ ถอนเงินระบุจำนวนและถอนเงินโดยเลือกจากเมนูของระบบ

โดเมนการกรอกรหัสผ่านคือ กรอกจากหน้าจอ

นำโดเมนระดับแม่และลูกเหล่านี้จัดเก็บในฐานข้อมูล SQL โดยสร้างเป็นฐานข้อมูลใหม่



ภาพที่ 4: อนุกรมวิธานระบบเอทีเอ็ม

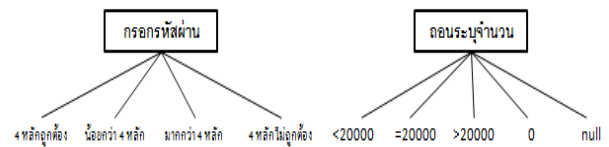
3.3 ใช้หลักการต้นไม้จำแนกข้อมูลออกแบบลักษณะอินพุต

ออกแบบลักษณะอินพุตที่ใช้สร้างกรณีทดสอบจัดเก็บภายใต้อนุกรมวิธานดังภาพที่ 5 เพื่อช่วยในการสร้างกรณีทดสอบแบบอัตโนมัติ ตัวอย่างการออกแบบลักษณะอินพุต

อินพุตของการกรอกรหัสผ่านเอทีเอ็มคือ ใส่รหัส 4 หลัก และเป็นรหัสถูกต้อง, ใส่น้อยกว่า 4 หลัก, ใส่มากกว่า 4 หลัก, ใส่ 4 หลักแต่รหัสไม่ถูกต้อง

อินพุตของการถอนระบุจำนวนคือ น้อยกว่า 20000, เท่ากับ 20000, มากกว่า 20000, 0, null

โดยอินพุตเหล่านี้ถูกกำหนดโดยผู้เกี่ยวข้องกับระบบเอทีเอ็มคือ ผู้ใช้หรือลูกค้า นักวิเคราะห์ระบบ นักทดสอบ และจะนำไปจัดเก็บในฐานข้อมูล SQL ที่ถูกสร้างใหม่เพื่อใช้ในระบบนี้ (ซึ่งจากตัวอย่างอินพุตที่ใส่ข้างต้นจะให้หลักการวิเคราะห์ขอบ และกรณีข้อมูลที่เป็น Invalid Data)



ภาพที่ 5: การจัดจำแนกอินพุตที่เป็นไปได้

3.4 ใช้กฎพึ่งพาอาศัยกันประยุกต์ใช้กับแนวทางที่ออกแบบ

กฎการพึ่งพาอาศัยกันจะถูกกำหนดเข้ากับต้นไม้จำแนกข้อมูลเพื่อช่วยลดทอนจำนวนกรณีทดสอบที่ไม่จำเป็นออกไปโดยใช้วิธีการ NOT ซึ่งได้อธิบายใน 2.3 กำหนดส่วน Exception ของยูสเคสเดสคริปชันช่วยในการลดทอนกรณีทดสอบดังภาพที่ 6 ซึ่งได้แสดงตัวอย่างอินพุตการกรอกรหัสผ่านและจำนวนเงิน ที่ถูกลดทอนโดย Exception

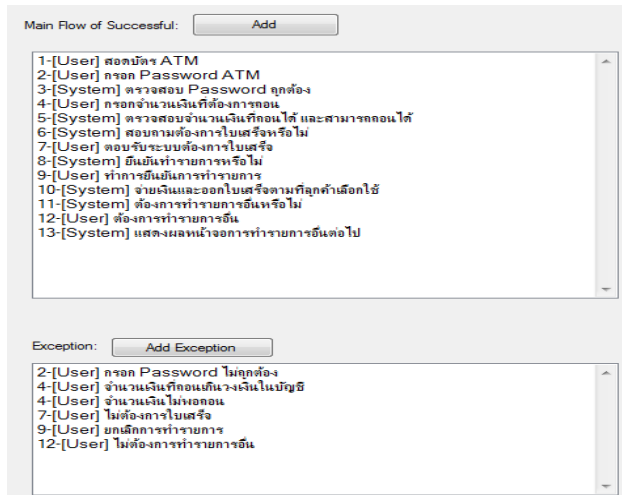
สำหรับการพัฒนาระบบงานวิจัยนี้จะมีตัวแปรที่มากขึ้นและจะแสดงการวัดผลใน 3.7

ลำดับ	4 หลัก ถูกต้อง	น้อยกว่า 4 หลัก	มากกว่า 4 หลัก	จำนวนเงิน ถูกต้อง	<20000	=20000	>20000	0	null
1	X				X				
2	X					X			
3	X						X		
4	X							X	
5	X								X
6		X			X				
7		X				X			
8		X					X		
9		X						X	
10		X							X
11			X		X				
12			X			X			
13			X				X		
14			X					X	
15			X						X
16				X	X				
17				X		X			
18				X			X		
19				X				X	
20				X					X

ภาพที่ 6: กรณีทดสอบที่ต้องใช้กฎเข้ามาช่วย

3.5 พัฒนาระบบการเขียนยูสเคสเดสคริปชัน

พัฒนาระบบในการเขียนยูสเคสเดสคริปชันเพื่อให้ผู้ใช้ระบบได้ทำการกรอกข้อมูลเกี่ยวกับยูสเคสนั้นๆ โดยใช้ ออกแบบหน้าจอดังภาพที่ 7 ระบบนี้จะทำการรับข้อมูลยูสเคสเดสคริปชันจากผู้ใช้และสามารถบันทึกข้อมูลสู่ฐานข้อมูล



ภาพที่ 7: หน้าจอระบบที่ใช้สร้างยูสเคสเดสคริปชัน

3.6 พัฒนาระบบในการสร้างกรณีทดสอบแบบอัตโนมัติ

ทำการสร้างระบบต่อยอดจากระบบเขียนยูสเคสเดสคริปชัน โดยสร้างทางเดินของฟังก์ชันการทำงานที่เกิดจากยูสเคสเดสคริปชันเพื่อช่วยตรวจสอบได้ว่าการสร้างกรณีทดสอบครอบคลุมสำหรับในทุกๆทางเดินของฟังก์ชัน จากนั้นทำการตรวจสอบซินนาริโอของยูสเคสเดสคริปชันกับฐานข้อมูล เพื่อทำการสร้างกรณีทดสอบอัตโนมัติจากหลักการต้นไม้จำแนกข้อมูลและอนุกรมวิธาน

กล่าวโดยละเอียดคือระบบจะนำทางเดินของยูสเคสเดสคริปชันมาสร้างเป็นกราฟทางเดินสำหรับอธิบายภาพการเดินทางของฟังก์ชัน โดยทำการคัดกรองกระทำของผู้ใช้มาจัดเรียงและสร้างเป็นกราฟดังภาพที่ 8 อีกทั้งกราฟนี้จะเป็นประโยชน์ต่อระบบด้วยการเป็นเครื่องช่วยตรวจสอบความครอบคลุมของกรณีทดสอบที่จะถูกสร้างขึ้นในขั้นตอนถัดไป

ขั้นตอนต่อมาคือทำการตรวจสอบทางเดินในยูสเคสเดสคริปชันแต่ละขั้นตอนว่าอยู่ภายใต้อนุกรมวิธานใด ทำการดึงข้อมูลอินพุตของต้นไม้จำแนกข้อมูลภายใต้อนุกรมวิธานนั้นๆและนำมาสร้างกรณีทดสอบ ข้อมูลจากภาพที่ 9 ระบบทำการสร้างกรณีทดสอบอัตโนมัติแบบครอบคลุมสูงสุดและทำ

การลดทอนกรณีทดสอบด้วย Exception ของเหตุการณ์จากยูสเคสเดสคริปชัน ได้ผลดังภาพที่ 9



ภาพที่ 8: กราฟทางเดินฟังก์ชันแสดงการกระทำของผู้ใช้ระบบ

จำนวน	4 นก	น้อยกว่า 4 นก	มากกว่า 4 นก	รถผ่าน ไม่ถูกต้อง	<-20000	-20000	>20000	0	null
1	X				X				
2	X					X			
3	X						X		
4	X							X	
5	X								X
6		X							
7			X						
8				X					

ภาพที่ 9: กรณีทดสอบอัตโนมัติถูกสร้างจากโปรแกรม

3.7 การทดสอบให้ครอบคลุมและเหมาะสม

จากการจัดจำพวกและแบ่งอินพุตให้เป็นกลุ่ม และทำขั้นตอนนี้อย่างรีเคอซีฟ แล้วสร้างตารางคอมบินชันจากต้นไม้จำแนกข้อมูล ซึ่งได้ชุดกรณีทดสอบออกมา ซึ่งถ้าจะทดสอบโดยความครอบคลุมสูงสุด ก็จะได้ชุดกรณีจำนวนมาก ซึ่งจะมีชุดกรณีทดสอบบางชุดที่เป็นไปไม่ได้ (Infeasible) หรือไม่จำเป็น ส่วนการทดสอบความครอบคลุมต่ำสุด ซึ่งจำนวนกรณีทดสอบจะเท่ากับจำนวนอินพุตที่มีค่ามากที่สุดโดเมน ซึ่งถ้าใช้การครอบคลุมต่ำสุดอาจจะพลาดกรณีทดสอบที่สำคัญไป ดังนั้นจึงใช้วิธีสร้างกรณีทดสอบให้ครอบคลุมมากที่สุดก่อน และทำการลดจำนวนกรณีทดสอบลง เพราะเป็นไปไม่ได้ที่จะทดสอบทุกกรณี ซึ่งจากการสร้างกราฟทางเดินของยูสเคสที่ได้รวมกรณีมี Exception จะทำการทดสอบที่ครอบคลุมทุกซินนาริโอที่เป็นไปได้ ซึ่งทางเดินที่มี Exception ก็จะตรงกับกฎการฟังพาทาศัย NOT ซึ่งจะทำให้ลดกรณีทดสอบได้จำนวนมาก ซึ่งจะควบคุมความเสี่ยงที่จะไม่ละทิ้งการทดสอบที่สำคัญไป เนื่องจากเราสร้างกรณีทดสอบโดยใช้ซินนาริโอทางเดินหลัก และ Exception ในยูสเคสเดสคริปชัน ส่วนความสัมพันธ์ OR จะเป็นไปโดยปริยาย

จากภาพที่ 9 มีการกรองรหัสผ่าน 4 ค่า กรอกจำนวนเงิน 5 กรณีที่เป็นไปได้ (ซึ่งใส่ค่าตามการวิเคราะห์ขอบของข้อมูล) การทดสอบระบบที่เริ่มที่ใช้ในงานวิจัย ทำการเพิ่มตัวแปรอีก 2 ตัวแปร คือต้องการใบเสร็จหรือไม่ และต้องการทำรายการธุรกรรมอื่นหรือไม่ ซึ่งหากทำการพิจารณาจำนวนอินพุตทั้งหมด จำนวนกรณีทดสอบจะเกิดจากคอมบินชัน $4*5*2*2$ เท่ากับ 80 กรณีทดสอบและสามารถใช้ Exception ในการลดทอนได้ผลดังภาพที่ 10 ซึ่งจำนวนกรณีทดสอบลดลง 71.25 เปอร์เซ็นต์ โดยกรณีทดสอบถูกสร้างขึ้นจากยูสเคสเดสคริปชันซึ่งอธิบายการทำงานในรูปแบบซินนารีโอ ทำให้แน่ใจได้ว่ากรณีทดสอบที่ถูกลดทอนนั้นเป็นกรณีทดสอบที่ถูกละทิ้งได้ และกรณีทดสอบที่ถูกสร้างมีพื้นฐานจากทางเดินหลักของระบบซึ่งเป็นกรณีทดสอบที่ควรถูกทดสอบอย่างแท้จริง อีกทั้งยังมีความน่าจะเป็นที่ผู้ใช้จะใช้งานตามทางเดินนี้ได้มากกว่าทางเดินอื่นๆ โดยไม่จำเป็นต้องทำการให้ค่าน้ำหนักความน่าจะเป็นเหมือนกับงานวิจัย [11] และผลลัพธ์จำนวนกรณีทดสอบยังมีความครอบคลุมเหมาะสมเป็นไปตาม [8]

วิธีการสร้างกรณีทดสอบ	จำนวนกรณีทดสอบ	จำนวนที่ลดลงจากครอบคลุมมากที่สุด	จำนวนที่ลดลง (%)
ครอบคลุมมากที่สุด	80	0	0
ครอบคลุมมากที่สุด + Exception	23	57	71.25
ครอบคลุมน้อยที่สุด	5	75	93.75
ครอบคลุมน้อยที่สุด + Dependency	8	72	90
Pairwise Testing	20	60	75

ภาพที่ 10: ตารางเปรียบเทียบจำนวนกรณีทดสอบโดยใช้แนวทางต่างๆ ของต้นไม้จำแนกข้อมูล

4. บทสรุป

การสร้างกรณีทดสอบให้ครอบคลุมถึงความต้องการและประหยัดเวลาในการทำงานยังคงเป็นปัญหาที่เกิดขึ้นบ่อยครั้งในการพัฒนาระบบคอมพิวเตอร์ บทความนี้ได้นำเสนอแนวคิดที่ช่วยในการสร้างกรณีทดสอบแบบอัตโนมัติเพื่อแก้ปัญหา โดยใช้หลักการอนุกรมวิธานและต้นไม้จำแนกข้อมูล วิธีการนี้มีข้อดีคือประหยัดเวลาในการทำการทดสอบแบบถดถอยและยังผนวกเข้ากับขั้นตอนการออกแบบระบบคือการสร้างยูสเคสเดสคริปชัน จุดประสงค์เพื่อที่จะสามารถแก้ไขงานออกแบบและสร้างกรณีทดสอบให้สอดคล้องซึ่งกันและกัน

โดยงานวิจัยประเภทนี้ที่ใช้ยูสเคสเดสคริปชันซึ่งอธิบายซินนารีโอมาช่วยสร้างต้นไม้จำแนกข้อมูลซึ่งเป็นการสร้างกรณีทดสอบในเชิงข้อมูลซึ่งยังไม่มีผู้ทำการวิจัยมาก่อน จึงยังไม่มี การเปรียบเทียบกับวิธีการอื่นมากนัก อย่างไรก็ตามระบบที่ถูกพัฒนาจากงานวิจัยนี้ยังมีข้อบกพร่อง ทั้งในเรื่องการเสีย ระยะเวลาการออกแบบในครั้งแรกของการติดตั้งระบบ หน้าจอการทำงานที่ยังใช้งานได้ยากในบางขั้นตอน ดังนั้นในอนาคต ผู้วิจัยจะได้พัฒนาระบบให้มีประสิทธิภาพมากยิ่งขึ้น

เอกสารอ้างอิง

- [1] P. M. Kruse, Joachim Wegener, "Test Sequence Generation from Classification Trees" *EU grant ICT-257574 (FITTEST)*.
- [2] P. M. Kruse, Joachim Wegener, "Test Sequence Generation from Classification Trees" *IEEE Fifth International Conference on Software Testing, Verification and Validation, 2012*
- [3] สิริมา นาคเต็ย, อำนาจ เปาะทอง, "กรอบแนวคิดการสร้างกรณีทดสอบจากคุณลักษณะความต้องการในรูปแบบฟอร์ม" *ACTIS, ISSN: 1906-9006, 2012.*
- [4] O. Taipale, K. Smolander, "Improving Software Testing by Observing Practice" *International Symposium on Empirical Software, 2006.*
- [5] J. A. Whittaker, "What Is Software Testing? And Why Is It So Hard?" *IEEE, 2000*
- [6] Wikipedia.org, "Classification Tree Method"
- [7] P. M. Kruse, Kiran Lakhota, "Multi Objective Algorithms for Automated Generation of Combinatorial Test Cases with the Classification Tree Method", *EU grant ICT-257574 (FITTEST)*.
- [8] E. Lehmann, J. Wegener "Test Case Design by Means of the CTE XL" *Research and Technology Alt-Moabit 96 a D-10559 Berlin*
- [9] B. Ramadoss, P. Prema, S. R. Balasundaram "Improving Test Coverage Using Merge Classification Tree" *International Conference and Workshop on Emerging Trends in Technology (ICWET 2010) – TCET, Mumbai, India.*
- [10] A. Krupp, W. Muller "A Systematic Approach to the Test of Combined HW/SW Systems" *EDAA 2010*
- [11] P. M. Kruse, M. Luniak "Automated Test Case using Classification Trees" *ASQ 2010*

โครงสร้างต้นไม้สำหรับแจกแจงความสัมพันธ์ระหว่างโมดูล Tree Structure for Association Identification between Modules

สกรณ์ บุษบง (Sakorn Bussabong)¹ และพิชโยทัย มัทธนาภิวัฒน์(Noppamas Pukkhem)²

สาขาวิศวกรรมคอมพิวเตอร์ มหาวิทยาลัยเทคโนโลยีสุรนารี จนครราชสีมา.

zbzagon@gmail.com¹, pmh@sut.ac.th²

บทคัดย่อ

ในปัจจุบันการเขียนโปรแกรมเชิงวัตถุเป็นที่นิยมอย่างมาก ในหมู่นักพัฒนาโปรแกรมทุกระดับ จากลักษณะเฉพาะของการเขียนโปรแกรมเชิงวัตถุทำให้การเปลี่ยนแปลง แก้ไขและการนำกลับมาใช้ใหม่ได้ง่าย ทำให้การเขียนโปรแกรมมีความสะดวกรวดเร็วมากขึ้น แต่ถ้าโปรแกรมที่ทำการเขียนนั้นมีความซับซ้อนมาก ทำให้มีคลาสเป็นจำนวนมาก ความสัมพันธ์ระหว่างคลาสและโมดูลก็มากตามไปด้วย เมื่อมีความจำเป็นที่จะต้องตรวจสอบโครงสร้างความสัมพันธ์ระหว่างโมดูลซึ่งอยู่ในคลาสหรือทดสอบการทำงาน ก็จะทำให้เกิดความยุ่งยากอันเนื่องมาจากความซับซ้อนของตัวโปรแกรมเอง

จากปัญหานี้ผู้วิจัยจึงได้ออกแบบโครงสร้างต้นไม้เพื่อใช้ในการระบุความสัมพันธ์ระหว่างโมดูล ในแต่ละคลาส โดยการปรับปรุง Abstract Syntax Tree ให้เหมาะสมในการแจกแจงโครงสร้างความสัมพันธ์เพื่อใช้แก้ปัญหาข้างต้นและเป็นต้นแบบในการพัฒนาต่อไป

คำสำคัญ: การเขียนโปรแกรมเชิงวัตถุ, การทดสอบ, โครงสร้างไวยากรณ์ต้นไม้, ความสัมพันธ์ของโมดูล

Abstract

Nowadays, object-oriented programming is very popular among developers of all levels.

Characteristic of object-oriented programming are to make the transition editing and reusability easier, make programming a lot more convenient. But if that program is very complex, containing a lot of classes and a lot of

relationships between classes and modules. It is necessary to examine the structure of relationships between these. In order to verify accuracy, or functional testing, it would be difficult because of the complexity of the program.

By this problem researchers have designed a tree structure to identify relationships between modules in each class by improving the Abstract Syntax Tree for proper identification of the structural relationship to the above solution and a prototype for further development.

Keywords: Object-oriented Programming, Testing, Abstract Syntax Trees, Module Relation

1. บทนำ

การเขียนโปรแกรมเชิงวัตถุเป็นที่นิยมอย่างมากในหมู่นักพัฒนาทุกระดับ ตั้งแต่ผู้เริ่มต้นไปจนถึงผู้เชี่ยวชาญ ในการพัฒนาโปรแกรมนั้นมักจะมี ความซับซ้อนมาก เนื่องจากความก้าวหน้าทางเทคโนโลยีทำให้การออกแบบโครงสร้างของโปรแกรมมีความซับซ้อนตามไปด้วย ดังนั้นการตรวจสอบว่าโปรแกรมที่พัฒนาขึ้นนั้นมีโครงสร้างความสัมพันธ์ระหว่างโมดูลถูกต้องตามการออกแบบจึงทำได้ยากมากขึ้น

ในการเขียนโปรแกรมเชิงวัตถุ นั้นมักจะมีคลาสจำนวนมาก ภายในคลาสนั้นก็จะประกอบไปด้วยโมดูลต่างๆมากมายซึ่งแต่ละโมดูลจะมีความสัมพันธ์ระหว่างโมดูลในคลาสดียวกัน หรือโมดูลในคลาสอื่นๆทำให้เกิดความสัมพันธ์ที่ซับซ้อนกันหลายชั้น

จุดนี้มักจะเป็นปัญหาในการตรวจสอบความถูกต้องของโปรแกรมเพื่อให้ตรงกับกรอกแบบ

ในเอกสารงานวิจัยนี้ จะนำเสนอเทคนิคและวิธีการที่ใช้ในการตรวจสอบความสัมพันธ์ระหว่างโมดูลโดยใช้ การปรับปรุง Abstract Syntax Tree ในการแจกแจงโครงสร้างของซอร์สโค้ดเพื่อระบุความสัมพันธ์ระหว่างโมดูลออกมาในรูปแบบของโครงสร้างต้นไม้

2. วรรณกรรมที่เกี่ยวข้อง

Abstract Syntax Tree (AST) เป็นโครงสร้างพื้นฐานที่ทำหน้าที่แสดงโครงสร้างของข้อมูลและซอร์สโค้ด ถูกนำไปประยุกต์ใช้กันอย่างแพร่หลาย เช่น Compiler, Interpreter รวมถึงการอธิบายเงื่อนไขต่างๆ เป็นต้น

Somka-ne Balla และ P. Mahatthanapiwat[1] ได้นำเสนองานวิจัยเกี่ยวกับความเป็นไปได้ของการนำ AST ใช้ในการแจกแจงโครงสร้างของซอร์สโค้ด เพื่อใช้ในการตรวจจับความแตกต่างของซอร์สโค้ด ซึ่งให้ผู้ใช้วิจัยได้แนวคิดในการใช้ AST แจกแจงโครงสร้างในระดับโมดูล งานวิจัยของ Junfei Huang และ Yunzhan Gong[2] ได้นำเสนอการทดสอบ BPEL (Web Services Business Process Execution Language) โดยใช้ต้นไม้ EMF ซึ่งมีความคล้ายคลึงกับ AST เพื่อใช้ในการอธิบายโครงสร้างของ BPEL ซึ่งมีแนวคิดเดียวกับงานวิจัยนี้คือมีการใช้ AST แจกแจงโครงสร้างเพื่อใช้ในการทดสอบ

Ira D. Baxter, Andrew Yahin, Leonardo Moura, Marcelo Sant'Anna และ Lorraine Bier[3] ได้นำเสนองานวิจัยที่เกี่ยวกับวิธีการตรวจจับการทำซ้ำของซอร์สโค้ด โดยใช้ AST มาประยุกต์ใช้ในการแจกแจงโครงสร้างของซอร์สโค้ด

ยังมีงานวิจัยของ Tahira Khatoun Priyansha Singh และ Shikha Shukla[4] ที่ได้นำเสนองานวิจัยเกี่ยวกับการตรวจจับการซ้ำกันของซอร์สโค้ดเช่นเดียวกันแต่มีการนำไปใช้กับภาษาจาวา ทำให้ผู้ใช้วิจัยเกิดแนวคิดในการนำ AST มาประยุกต์ร่วมกับภาษาเชิงวัตถุ

งานวิจัยของ Guo Tao, Dong Guowei, Qin Hu,

Cui Baojiang[5] เป็นการพัฒนาอัลกอริทึมเพื่อตรวจจับการถูกปลอมแปลงซอร์สโค้ด โดยเฉพาะ ภายในคำสั่ง if และ while โดยใช้พื้นฐานของ AST ซึ่งมีการคำนวณ Hash Value ของทุก Node และนำค่า Hash Value ของแต่ละ Node มาเปรียบเทียบกับ

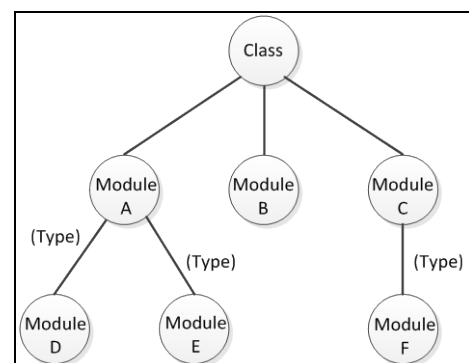
งานวิจัยของ Iulian Neamtii, Jeffrey S. Foster และ Michael Hicks[6] ได้นำเสนองานวิจัยเกี่ยวกับการตรวจสอบการพัฒนาซอฟต์แวร์ภาษาซี โดยใช้เทคนิคการจับคู่ของ AST เน้นไปที่ชนิดและตัวแปรเพื่อตรวจสอบความแตกต่างของเวอร์ชันของซอฟต์แวร์ งานวิจัยของ G. Fischer, J. Lusiardi และ J. Wolff von Gudenberg[7] ได้นำเสนอ เกี่ยวกับบทบาทของ AST ต่อการพัฒนาซอฟต์แวร์

จากงานวิจัยเหล่านี้ทำให้เกิดแนวคิดที่จะนำ AST มาประยุกต์ใช้ในการแจกแจงโครงสร้างของซอร์สโค้ด เพื่อสร้างต้นไม้สำหรับแจกแจงความสัมพันธ์ระหว่างโมดูล เพื่อใช้ในการตรวจสอบความสัมพันธ์ระหว่างโมดูลต่อไป

3. วิธีการดำเนินการวิจัย

งานวิจัยนี้ได้มีการปรับปรุง AST ขึ้นมาเพื่อมุ่งเน้นไปที่การแจกแจงความสัมพันธ์ระหว่างโมดูล

AST นั้นมีโครงสร้างเหมือนต้นไม้ทั่วไปเดิมที่ AST นั้นได้ถูกนำไปประยุกต์ได้หลากหลาย หนึ่งในนั้นคือทำหน้าที่แสดงโครงสร้างของซอร์สโค้ด และแจกแจงในรูปแบบของต้นไม้



ภาพที่ 1: โครงสร้าง AST ที่ได้ปรับปรุง

โครงสร้างต้นไม้ ที่ได้ปรับปรุงขึ้นมานั้น แต่ละต้นไม้จะเป็นตัวแทนของแต่ละคลาส ดังภาพที่ 1 ต้นไม้เหล่านี้จะอยู่รวมกันเป็นกลุ่ม โดยมีการเชื่อมโยงกันระหว่างต้นไม้

โครงสร้างต้นไม้ที่ผู้วิจัยได้ปรับปรุงขึ้นมานั้นประกอบไปด้วย โครงสร้างหลัก 3 ส่วน ดังต่อไปนี้

Root เป็นตัวแทนของชื่อคลาส กำหนดค่าภายในคือ ชื่อของคลาส

Node เป็นตัวแทนของแต่ละโมดูล โดยจะเก็บชื่อของโมดูลที่อยู่ภายในคลาสและพารามิเตอร์ของโมดูล

Edge เป็นตัวแสดงความสัมพันธ์ระหว่าง Node จะเก็บค่าอาร์กิวเมนต์ของโมดูลที่ส่งให้กันรวมถึงโมดูลที่ไม่มีการส่งอาร์กิวเมนต์ให้กันด้วย

โครงสร้างทั้ง 3 ส่วนนี้จะมี Attribute อยู่ในซึ่งใช้ในการระบุตัวตนและแสดงความสัมพันธ์ต่างๆ

Node แต่ละ Node จะเรียงตามการเรียกใช้ Node อื่นๆ โดยมี Edge แสดงความสัมพันธ์ระหว่าง Node โดยที่ Edge จะมีค่าอาร์กิวเมนต์ของโมดูลผู้เรียกซึ่งจะสัมพันธ์กับพารามิเตอร์ ของโมดูลผู้ถูกเรียก Edge ที่เชื่อมระหว่าง Root และ Child นั้นจะไม่มี Attribute ภายใน เนื่องจากไม่ได้เกิดความสัมพันธ์ระหว่างโมดูล เพียงแต่เป็นการแสดงโครงสร้างต้นไม้ให้สมบูรณ์

ระดับของ Node ภายในต้นไม้จะแสดงถึงลำดับการเรียกใช้โมดูล โดยที่ Node ที่มีการเรียกโมดูลอื่นๆ ก็จะมี Node มาเชื่อมต่อจาก Node นี้ เป็นลำดับชั้นลงไปเรื่อยๆ ซึ่งสุดท้ายแล้ว Node ที่ไม่ได้ถูกเรียกใช้จะอยู่ระดับบนสุด และ Node ที่ไม่ได้เรียกโมดูลใดๆจะอยู่ระดับล่างสุด

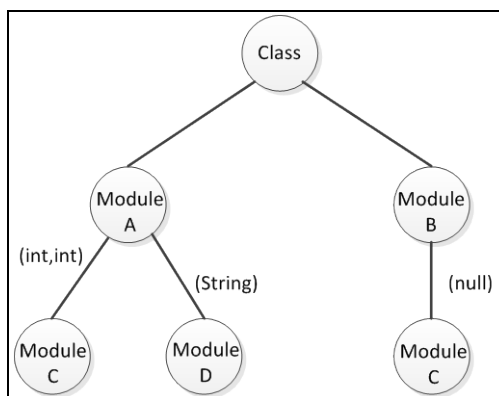
เนื่องจากกระบวนการนี้อยู่ภายหลังกระบวนการพัฒนาซอฟต์แวร์ ซอร์สโค้ดที่นำเข้ามาในกระบวนการนี้จึงผ่านการ Compile จาก Compiler เรียบร้อยแล้ว ทำให้ไม่จำเป็นต้องคำนึงถึงการ Encapsulation เนื่องจาก Compiler ทำหน้าที่ตรวจสอบความถูกต้องทั่วไปให้แล้ว

ในกรณีที่โมดูลมีการ Overloading คือการมีโมดูลที่มีชื่อเหมือนกันแต่มีพารามิเตอร์ต่างกัน จะมีการแสดงโมดูลแยกออกจากกันตามโครงสร้างของการเรียกใช้งานตามภาพที่ 2 ซึ่ง โมดูล

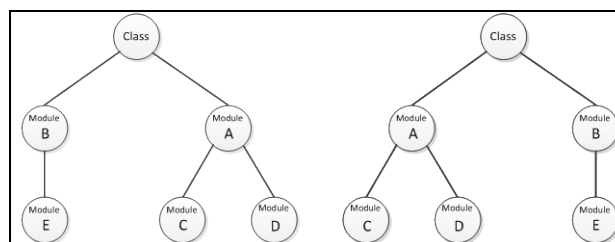
C จะถูกเรียกใช้ทั้ง 2 โมดูลแต่อาร์กิวเมนต์ที่ได้รับจะแตกต่างกันไปตามพารามิเตอร์ของโมดูลนั้นๆ

จำนวนของต้นไม้ที่ขึ้นอยู่กับการปรับจำนวนของคลาส โดยจะไม่สนใจคลาสที่เป็น Interface เนื่องจากโมดูลภายใน Interface นั้นไม่มีการ Implement ทำให้ไม่มีผลต่อโครงสร้างต้นไม้สำหรับแจกแจงความสัมพันธ์นี้

รูปร่างของต้นไม้แต่ละต้นนั้นจะแปรผันตามลำดับ และความซับซ้อนของโมดูลที่มีความสัมพันธ์กันทำให้การจัดวางตำแหน่งของโมดูล การจัดลำดับของ Node ของโมดูลระดับแรกนั้น สามารถสลับตำแหน่งได้ขึ้นอยู่กับกระบวนการระบุโมดูล ดังนั้น ซอร์สโค้ดที่นำเข้าสู่กระบวนการสร้างสามารถสร้างต้นไม้ได้หลายแบบ แต่สามารถให้อธิบายโครงสร้างซอร์สโค้ดได้เช่นเดียวกันดังภาพที่ 3



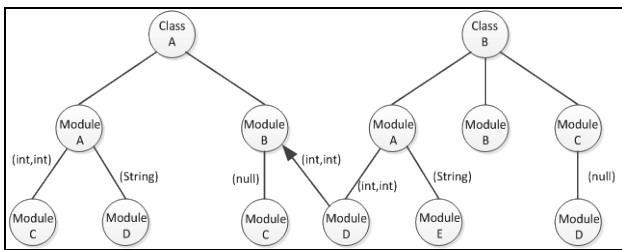
ภาพที่ 2: ภาพโครงสร้าง AST ที่ปรับปรุงแล้วกับโมดูล Overloading



ภาพที่ 3: ภาพความแตกต่างของโครงสร้างต้นไม้จากซอร์สโค้ดเดียวกัน

การเรียกใช้โมดูลจากคลาสไปยังโมดูลจากคลาสนั้นจะทำให้เกิดโครงสร้างความสัมพันธ์ระหว่างต้นไม้โดยมี Edge เชื่อมจาก Node ของต้นไม้ต้นหนึ่งไปยังอีกต้นหนึ่ง โดยที่ความสัมพันธ์ของต้นไม้จะขึ้นอยู่กับความสัมพันธ์ระหว่างโมดูล ถ้าแต่ละคลาสเกิดความสัมพันธ์ระหว่างโมดูลจำนวนมากต้นไม้ก็จะมี Edge เชื่อมต่อกันระหว่างต้นไม้เป็นจำนวนมากตามไปด้วย

Edge ที่เชื่อมระหว่างโมดูลที่อยู่ภายในคลาสที่ต่างกันนั้นจำเป็นที่จะต้องมีการกำหนดทิศทางอย่างชัดเจน เนื่องจากการแสดงโมเดลโครงสร้างต้นไม้ในแต่ละต้นไม้เป็นอิสระต่อกันถ้าไม่มีการกำหนดทิศทางของ Edge อย่างชัดเจนจะทำให้เกิดความสับสนระหว่างโมดูลผู้เรียกและผู้ถูกเรียกโดยมี Edge เชื่อมโยงดังภาพที่ 4 ซึ่งเป็นภาพที่แสดงความสัมพันธ์ระหว่างต้นไม้ โดยที่ความสัมพันธ์ข้ามคลาสนั้นเกิดขึ้นที่ โมดูล D ภายในคลาส B และ โมดูล B ภายในคลาส A



ภาพที่ 4: ภาพความสัมพันธ์ระหว่างต้นไม้

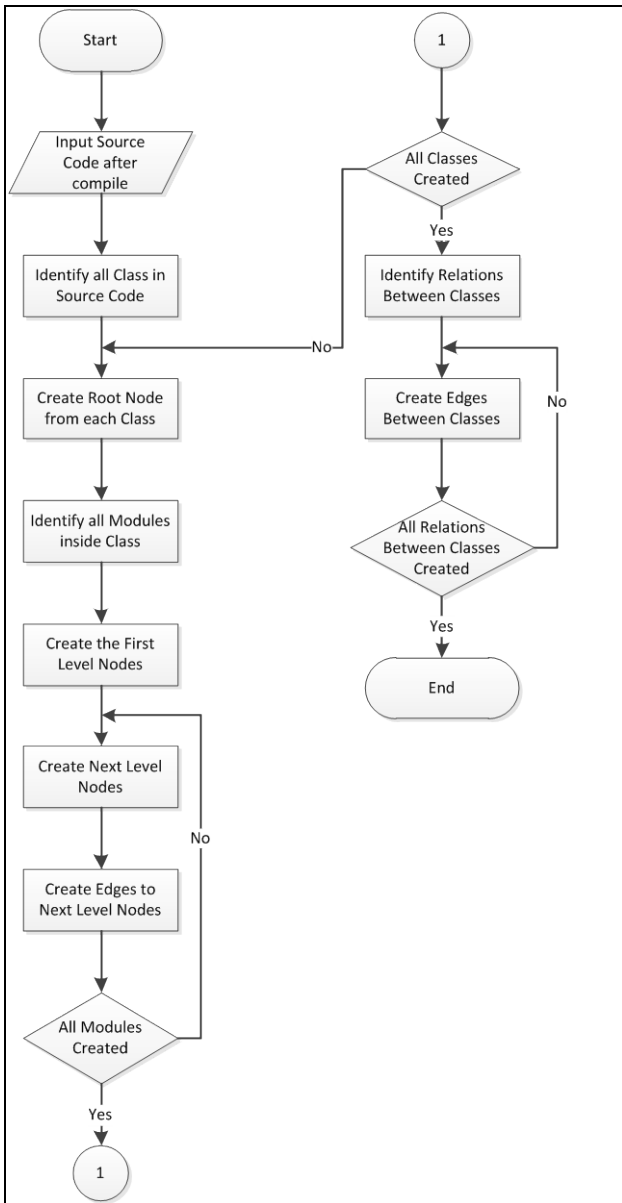
ในกระบวนการสร้างโครงสร้างต้นไม้สำหรับแจกแจงความสัมพันธ์ระหว่างโมดูลนั้นมีขั้นตอนดัง ภาพที่ 5 ซึ่งมีรายละเอียดดังนี้

1. Input Source Code after Compile : เป็นการนำเข้าซอร์สโค้ดที่ผ่านการ Compile แล้ว
2. Identify All Class in Source Code : เป็นการระบุคลาสทั้งหมดที่มีในซอร์สโค้ด รวมถึงการตรวจสอบว่าเป็น Interface หรือไม่
3. Create Root Node from each Class : สร้าง Root โดยกำหนด Attribute เป็น ชื่อของคลาส

4. Identify all Modules inside Class : นำข้อมูลของคลาสมาจำแนกโมดูลภายในโดยเริ่มจากระบุโมดูลที่มีทั้งหมดภายในคลาส และแยกระหว่างโมดูลที่เป็นผู้เรียก ผู้ถูกเรียก และเป็นทั้งสองอย่าง รวมถึงการแจกแจงพารามิเตอร์ของแต่ละโมดูลด้วย
 5. Create the First Level Nodes : กำหนด Node ระดับแรกจากโมดูลที่ไม่มีผู้เรียกก่อน แล้วใช้โมดูลระดับแรกที่กำหนดไว้แล้วมาระบุโมดูลที่ถูกเรียกจากโมดูลระดับแรก
 6. Create Next Level Nodes : สร้าง Node ระดับถัดมาโดยอ้างอิงจาก Node ระดับก่อนหน้า จากนั้นตรวจสอบว่า Node นั้นมีการเรียกโมดูลหรือไม่ ถ้ามีการเรียกโมดูลอื่นจะนำโมดูลที่ถูกเรียกมาสร้างเป็นลำดับต่อไป ถ้าไม่มีการเรียกใช้โมดูลอื่น ก็จะข้ามโมดูลนี้ไป
 7. Create Edges to Next Level Nodes : กำหนด Edge จาก พารามิเตอร์ ของโมดูลที่ถูกเรียก ถ้าโมดูลระดับแรกนี้ไม่มีการเรียกโมดูลอื่นก็จะข้ามโมดูลนี้ไป
 8. All Modules Created : ทำขั้นตอนที่ 6 และ 7 ซ้ำจนครบทุกโมดูล
 9. All Classes Created : ทำขั้นตอนที่ 3 ถึง 8 จนครบทุกคลาส
 10. Identify Relations Between Classes : เป็นการระบุความสัมพันธ์ของแต่ละคลาส เพื่อกำหนดโมดูลที่จะเป็นจุดเชื่อมต่อระหว่างคลาส
 11. Create Edge Between Classes : สร้าง Edge เชื่อมโมดูลที่มีการเรียกข้ามคลาส พร้อมทั้งมีการกำหนดพารามิเตอร์ ของโมดูลที่ถูกเรียก
 12. All Relation Between Classes Created : ทำขั้นตอนที่ 11 ซ้ำจนครบทุกความสัมพันธ์ระหว่างโมดูลข้ามคลาสที่เกิดขึ้น
- หลังจากทำตามขั้นตอนดังกล่าวเสร็จสิ้นแล้วจะได้กลุ่มของโครงสร้างต้นไม้ที่มีความสัมพันธ์กัน ซึ่งกลุ่มของต้นไม้ที่ได้มานั้นเป็นโครงสร้างต้นไม้ที่สามารถแสดงความสัมพันธ์ระหว่าง

โมดูลภายในคลาสเดียวกันและความสัมพันธ์ระหว่างโมดูลภายในคลาสที่มีความสัมพันธ์กับคลาสอื่นๆ

ทั้งหมด 4 คลาส มีโมดูลรวมกันทั้งหมด 18 โมดูล และมีความสัมพันธ์ระหว่างโมดูลข้ามคลาสทั้งหมด 3 แห่ง นำซอร์สโค้ดเหล่านี้มาเข้าสู่กระบวนการสร้างโครงสร้างต้นไม้



ภาพที่ 5: ภาพผังงาน(Flow chart) ของกระบวนการสร้างโครงสร้างต้นไม้

```

1 package sample;
2 public class Class_A {
3     public void ModuleAA() {
4         ModuleAD(5, 7);
5         ModuleAE(1f);
6         new Class_D().ModuleDB();
7     }
8     public void ModuleAB() {
9         new Class_B().ModuleBA("a");
10    }
11    public void ModuleAC() {
12        ModuleAF(Boolean.TRUE);
13    }
14    public int ModuleAD(int a, int b) {
15        return a + b;
16    }
17    public void ModuleAE(float a) {
18    }
19    public void ModuleAF(Boolean a) {
20    }
21 }

1 package sample;
2 public class Class_B {
3     public void ModuleBA(String a) {
4         ModuleBB(a, a);
5         ModuleBB(5.142);
6     }
7     public void ModuleBB(String a, String b) {
8         new Class_C().ModuleCA(5, b);
9     }
10    public void ModuleBB(double a) {
11    }
12 }

1 package sample;
2 public class Class_C {
3     public void ModuleCA(int a, String b) {
4         ModuleCC(a, b);
5         ModuleCD('a');
6     }
7     public void ModuleCB() {
8     }
9     public void ModuleCC(int a, String b) {
10        ModuleCE(a);
11        ModuleCF(true);
12    }
13    public void ModuleCD(char a) {
14    }
15    public void ModuleCE(int a) {
16    }
17    public void ModuleCF(Boolean a) {
18    }
19 }

1 package sample;
2 public class Class_D {
3     public void ModuleDA() {
4         ModuleDC(5.55);
5     }
6     public void ModuleDB() {
7         ModuleDC(5.55);
8     }
9     public void ModuleDC(double a) {
10    }
11 }
    
```

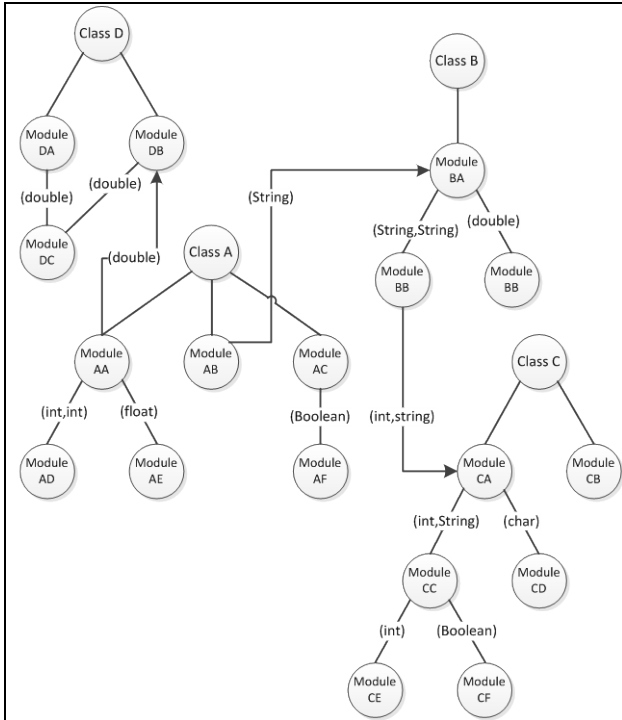
ภาพที่ 6: ภาพซอร์สโค้ดที่ใช้ในการทดสอบ

4. ผลการดำเนินงาน

จากขั้นตอนที่กล่าวมาผู้วิจัยได้ดำเนินการทดลองกระบวนการสร้างโครงสร้างต้นไม้โดยใช้ซอร์สโค้ดภาษาจาวาดังภาพที่ 6 ซอร์สโค้ดดังกล่าวประกอบไปด้วยคลาส

หลักจากที่กระบวนการสร้างต้นไม้สำหรับแจกแจงความสัมพันธ์ระหว่างโมดูลเสร็จสิ้นแล้ว จะได้กลุ่มของต้นไม้

ดังภาพที่ 7 ซึ่งจะเห็นได้ว่าโครงสร้างต้นไม้ไม่สามารถแจกแจงโครงสร้างของซอร์สโค้ด และสามารถนำไปใช้ตรวจสอบความสัมพันธ์ระหว่างโมดูลได้สะดวกเนื่องจากโครงสร้างมีความเรียบง่ายและง่ายต่อการทำความเข้าใจ



ภาพที่ 7: ภาพกลุ่มของโครงสร้างต้นไม้ที่จำลองขึ้นมาจากซอร์สโค้ดในภาพที่ 6

5. สรุป

โครงสร้างต้นไม้สำหรับแจกแจงความสัมพันธ์ระหว่างโมดูลเป็นโครงสร้างต้นไม้ที่จะช่วยให้การตรวจสอบความสัมพันธ์ระหว่างโมดูลนั้นง่ายขึ้น สามารถนำไปใช้ร่วมกับการบวนการทดสอบโปรแกรมได้และเป็นแนวคิดที่สามารถนำไปพัฒนาเครื่องมือสร้างโครงสร้างต้นไม้สำหรับแจกแจงความสัมพันธ์ระหว่างโมดูลแบบอัตโนมัติได้ อย่างไรก็ตามแนวคิดนี้ยังคงต้องพึ่งมนุษย์ในการจำลองโครงสร้าง อีกทั้งยังไม่ได้ถูกนำไปใช้จริงและในขั้นการทดสอบนั้นได้ใช้ซอร์สโค้ดที่มีความซับซ้อนไม่มากในการทดสอบ

ปัจจุบันผู้เขียนกำลังพัฒนาโครงสร้างต้นไม้สำหรับแจกแจงความสัมพันธ์ระหว่างโมดูลในภาษาจาวาให้เป็นเครื่องมือเพื่อนำไปใช้ร่วมกับการสร้าง Integration Test Case ต่อไป

เอกสารอ้างอิง

- [1] Somka-ne Balla, P. Mahatthanapiwat, "Difference Detection Using Abstract Syntax Tree", *The 23rd National Graduate Research Conference (GRAD23), Rajamangala University of Technology Isan, Nakorn Ratchasima, Thailand, Oct 23-24, 2011.*
- [2] Junfei Huang , Yunzhan Gong, "An EMF Activity Tree Based BPEL Defect Pattern Testing Method", *2nd International Conference on Computer Engineering and Technology*, vol. 7, pp. 468-471, 2010.
- [3] Ira D. Baxter, Andrew Yahin, Leonardo Moura, Marcelo Sant'Anna, Lorraine Bier, "Clone Detection Using Abstract Syntax Trees", *International Conference on Software Maintenance*, pp. 368-377, 1998.
- [4] Tahira Khatoon, Priyansha Singh, Shikha Shukla, "Abstract Syntax Tree Based Clone Detection for Java Projects", *IOSR Journal of Engineering*, Vol 2, pp. 45-47, 2012.
- [5] Guo Tao , Dong Guowei, Qin Hu, Cui Baojiang, "Improved Plagiarism Detection Algorithm Based on Abstract Syntax Tree", *2013 Fourth International Conference on Emerging Intelligent Data and Web Technologies*, pp. 714-719, 2013.
- [6] Iulian Neamtii, Jeffrey S. Foster, Michael Hicks, "Understanding Source Code Evolution Using Abstract Syntax Tree Matching", *MSR '05 Proceedings of the 2005 international workshop on Mining software repositories*, pp. 1-5, 2005.
- [7] G. Fischer, J. Lusiardi, J. Wolff von Gudenberg, "Abstract Syntax Trees – and their Role in Model Driven Software Development", *International Conference on Software Engineering Advances(ICSEA 2007)*, pp. 38, 2007.

การค้นหาโรงแรมโดยการประยุกต์ใช้หลักการทำเหมืองข้อมูลความคิดเห็น

An Application of Hotel Searching Based on Opinion Mining

จันสธา ศรีสุวรรณ (Janasthar Srisuan)¹ และอนันตพร หารรรษคุณาลัย (Anantaporn Hanskunatai)²

สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

¹s3650801@kmitl.ac.th, ²ksananta@kmitl.ac.th

บทคัดย่อ

บทความนี้นำเสนอการประยุกต์ใช้หลักการทำเหมืองข้อมูลความคิดเห็นในการค้นหาโรงแรมจากโรงแรมทั้งหมด 15 โรงแรมในจังหวัดภูเก็ต โดยคัดเลือกข้อมูลความคิดเห็นทั้งหมด 293 ความคิดเห็น เพื่อนำมาสร้างโมเดลสำหรับการค้นหาด้วยตัวจำแนกแบบเบย์อย่างง่าย ผลลัพธ์ที่ได้จะแสดงลำดับโรงแรมจากคำค้นหาด้วยการนำค่าความน่าจะเป็นที่ได้จากตัวจำแนกแบบเบย์อย่างง่าย ในการวัดประสิทธิภาพของโมเดลได้ทำการเปรียบเทียบวิธีการค้นหาแบบตรรกะและวิธีการค้นหาแบบบอยเยอร์-มัวร์ ซึ่งปรากฏว่าตัวจำแนกแบบเบย์อย่างง่ายให้ผลลัพธ์ในการจัดอันดับที่ดีกว่าวิธีค้นหาทั้งสอง นอกจากนี้โมเดลที่นำเสนอใช้เวลาในการค้นหาเร็วที่สุดเมื่อเทียบกับวิธีอื่น

คำสำคัญ: ตัวจำแนกแบบเบย์อย่างง่าย การค้นหาโรงแรม การทำเหมืองข้อมูลความคิดเห็น

Abstract

This paper presents an application of opinion mining for hotel searching where selected as a dataset 293 reviews of 15 hotels in Phuket. This dataset is used to construct the searching model based on Naïve Bayes classifier. The output of the application is a list of hotel name ranking by hotel probability related to user keywords. The performance of the proposed model was compared with two classical searching methods: Boolean searching and Boyer-Moore string search. The results show that the Naïve Bayes Classifier model provides the highest average Rank_accuracy. In addition, the

proposed model also takes the fastest time in searching when compared which the other techniques.

Keyword: Naïve Bayes classifier, hotel searching, opinion mining.

1. บทนำ

ปัจจุบันการท่องเที่ยวภายในประเทศไทยได้เจริญเติบโตเพิ่มขึ้น และอุตสาหกรรมโรงแรมที่พักต่างๆ เพิ่มขึ้นมากมายด้วยเช่นกัน เพื่อรองรับการแข่งขันของประชาคมเศรษฐกิจอาเซียน (AEC) ยิ่งทำให้เกิดการแข่งขันกันเพื่อดึงดูดนักท่องเที่ยวทั้งชาวไทยและต่างชาติ การประชาสัมพันธ์และโฆษณาจึงมีความสำคัญอย่างยิ่ง ซึ่งเว็บไซต์เป็นอีกช่องทางหนึ่งในการประชาสัมพันธ์ที่ดี ทั้งเว็บไซต์ของสถานที่ท่องเที่ยวเอง หรือเว็บไซต์รวบรวมข้อมูลของสถานที่ต่างๆ ไว้คอยบริการ จึงทำให้ข้อมูลบนอินเทอร์เน็ตมีเพิ่มขึ้นตามเป็นลำดับ จากข้อมูลที่มีอยู่มากมายบนเว็บไซต์ที่รวบรวมข้อมูลสถานที่ท่องเที่ยวต่างๆ ซึ่งเปิดให้ผู้ใช้งานเข้ามามีส่วนร่วมในการวิจารณ์ หรือติชมสถานที่นั้นๆ ทางผู้วิจัยจึงได้สังเกตเห็นถึงประโยชน์ของข้อมูลดังกล่าว ซึ่งเป็นข้อมูลที่มาจากผู้ใช้งานทั่วไปหรือผู้เชี่ยวชาญ ดังนั้นงานวิจัยนี้มีวัตถุประสงค์เพื่อทำการค้นหาโรงแรมและเรียงลำดับ โรงแรมที่ตรงกับความต้องการของผู้ใช้ จากการพิจารณาชุดคำค้นหา โดยประยุกต์ใช้หลักการทำเหมืองข้อมูลความคิดเห็น (Opinion mining) ซึ่งในงานวิจัยนี้ได้เลือกใช้เทคนิคตัวจำแนกแบบเบย์อย่างง่ายมาทำการสร้างโมเดลการค้นหาโรงแรม

โดยทั่วไปแล้วตัวจำแนกแบบเบย์อย่างง่ายเป็นที่นิยมในการนำมาใช้ทำเหมืองข้อมูลความคิดเห็น ตัวอย่างเช่น การจำแนกทัศนคติที่ดีหรือไม่ดีของหนังสือ [1] การวิเคราะห์สถานะบน

เฟสบุคว่ามีความคิดเห็นเป็นบวก ลบ หรือ ปกติ [2] การจำแนกทัศนคติของหนังสือ [3] การจำแนกความคิดเห็นจากการอ่านข่าวทางด้านสังคมและการเมืองเพื่อให้ตรงกับความต้องการของผู้อ่าน [4]

เนื้อหาในส่วนถัดไปของบทความนี้กล่าวถึงทฤษฎีในบทที่ 2 การจำแนกแบบเบย์อย่างง่ายรวมทั้งวิธีการค้นหาแบบตรรกะและวิธีการค้นหาแบบบอยเซอร์-มัวร์ หัวข้อที่ 3 กล่าวถึงขั้นตอนการดำเนินงานวิจัยตั้งแต่ขั้นตอนการเตรียมข้อมูล สร้างโมเดล และการนำโมเดล ไปใช้ หัวข้อที่ 4 อธิบายถึงการออกแบบการทดลองและแสดงผลการทดลอง หัวข้อที่ 5 สรุปผลการทดลองโดยแสดงประสิทธิภาพของโมเดลที่นำเสนอ

2. ทฤษฎีที่เกี่ยวข้อง

ในหัวข้อนี้จะกล่าวถึงทฤษฎีที่ใช้และงานวิจัยที่เกี่ยวข้อง คือ ตัวจำแนกแบบเบย์อย่างง่าย (Naïve-Bayes classifier) [5] วิธีการค้นหาแบบตรรกะ (Boolean searching) [6] และวิธีการค้นหาแบบบอยเซอร์-มัวร์ (Boyer-Moore string search) [6]

2.1 ตัวจำแนกแบบเบย์อย่างง่าย

ตัวจำแนกแบบเบย์อย่างง่ายอาศัยหลักทฤษฎีความน่าจะเป็น โดยการคำนวณค่าความน่าจะเป็นแบบมีเงื่อนไข กล่าวคือการหาความน่าจะเป็นของการเกิดเหตุการณ์หนึ่งโดยมีเงื่อนไขว่ามีอีกเหตุการณ์หนึ่งเกิดขึ้นอยู่ก่อนแล้ว เขียนเป็นสัญลักษณ์ทางคณิตศาสตร์ คือ $P(A|B)$ โดยที่ A และ B แทนเหตุการณ์ เช่น กำหนดให้ A แทนเหตุการณ์นักศึกษาสอบตกวิชาคอมพิวเตอร์ และ B แทนเหตุการณ์ที่นักศึกษาลเล่นเฟสบุค เพราะฉะนั้น $P(A|B)$ หมายถึงความน่าจะเป็นที่นักศึกษาสอบตกวิชาคอมพิวเตอร์เมื่อนักศึกษาลเล่นเฟสบุค โดยมีสูตรการคำนวณความน่าจะเป็นแบบมีเงื่อนไขดังสมการที่ (1)

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)} \quad (1)$$

การประยุกต์ใช้ทฤษฎีแบบเบย์กับการจำแนกข้อมูลสามารถปรับให้สอดคล้องกับการจำแนกข้อมูลได้ดังสมการที่ (2)

$$P(\text{class}|\text{attribute}) = \frac{P(\text{class})P(\text{attribute}|\text{class})}{P(\text{attribute})} \quad (2)$$

เมื่อ $P(\text{class}|\text{attribute})$ คือความน่าจะเป็นในการเกิด class เมื่อมี attribute เกิดขึ้น

$P(\text{class})$ คือความน่าจะเป็นในการเกิด class

$P(\text{attribute})$ คือความน่าจะเป็นในการเกิด attribute

เมื่อข้อมูลแต่ละตัวอย่างมีจำนวนคุณลักษณะ (attribute) มากกว่าหนึ่งตัวจะเขียนสมการที่ (2) ใหม่ได้ดังสมการที่ (3)

$$P(\text{class}|a_1, a_2 \dots a_n) = \frac{P(\text{class})P(a_1, a_2 \dots a_n|\text{class})}{P(a_1, a_2 \dots a_n)} \quad (3)$$

ในกรณีที่เกิดคุณลักษณะ a_i ใดๆ ไม่ขึ้นตรงต่อกันสามารถเปลี่ยนวิธีคำนวณ $P(a_1, a_2 \dots a_n|\text{class})$ ใหม่ได้ดังสมการที่ (4)

$$P(a_1, a_2 \dots a_n|\text{class}) = \prod_{i=1}^n P(a_i|\text{class})$$

$$P(a_1, a_2 \dots a_n|\text{class}) = P(a_1|\text{class})P(a_2|\text{class}) \dots P(a_n|\text{class}) \quad (4)$$

ดังนั้นการจำแนกแบบเบย์อย่างง่ายจึงสามารถทำได้โดยการหาค่าความน่าจะเป็น $P(a_1, a_2 \dots a_n|\text{class})$ โดยขึ้นอยู่กับประเภทข้อมูลที่มีการเรียนรู้ในชุดข้อมูลนั้น เมื่อ c_j เป็นสมาชิกของ Class ต่างๆ โดยในการจำแนกประเภทจะพิจารณาค่าความน่าจะเป็นที่มีค่ามากที่สุดได้ดังสมการที่ (5)

$$\text{Classify}(\text{data}) = \underset{c_j \in C}{\text{argmax}} P(c_j) \prod_{i=1}^n P(a_i|c_j) \quad (5)$$

2.2 วิธีการค้นหาแบบตรรกะ

ในการค้นหาแบบตรรกะนี้ จะใช้วิธีการสืบค้นเอกสารโดยสร้างดัชนีจากเอกสารและนำคำค้นหาเพื่อมาเปรียบเทียบกับดัชนีของเอกสาร โดยในการค้นหาอาจจะมีหลายๆคำค้นหาได้ ซึ่งขึ้นอยู่กับความต้องการของผู้ใช้ โดยในการค้นหาจะมีการสร้างประโยคที่เป็นเงื่อนไขสำหรับการค้นหาด้วยการใช้คำเชื่อมคำค้นหา คือ And, Or หรือ Not ซึ่งในการค้นหาจะให้ผลลัพธ์ออกมาว่า Match หรือ Not match กับเอกสารเท่านั้นดังตัวอย่างตารางที่ 1

ตารางที่ 1: ตัวอย่างเอกสารที่ต้องการค้นหา

Document	Text
1	Pease porridge hot, pease porridge cold.
2	Pease porridge in the pot.
3	Nine days old.
4	Some like it in the pot. Some like it cold
5	Some like it in
6	Nine days old.

จากตารางที่ 1 หากมีคำค้นหาตัวอย่าง 1 ชุด คือ (hot, cold) โดยใช้คำเชื่อมคำค้นหา And จะได้ผลลัพธ์คือ Document 1 แต่หากใช้คำเชื่อมคำค้นหา Or จะได้ผลลัพธ์คือ Document 1, 4

2.3 วิธีการค้นหาของบอยเยอร์-มัวร์

วิธีการค้นหาของบอยเยอร์-มัวร์นี้จะทำการเทียบตัวอักษรในคำค้นหา (pattern) กับตัวอักษรจากข้อความ (text) จากขวาไปซ้ายหากตัวอักษรที่ตำแหน่งเดียวกันมีค่าไม่เหมือนกัน (Miss-match) จะทำการเลื่อนตำแหน่งของ pattern โดยพิจารณาตัวอักษรที่ Miss-match ของ text แล้วหาตำแหน่งที่ใกล้ที่สุดของตัวอักษรนั้นใน pattern (Bad-character shift) และหากตัวอักษรไม่ปรากฏใน pattern จะทำการเลื่อนตำแหน่งไปเท่ากับจำนวนตัวอักษรทั้งหมดของ pattern หากการเลื่อนทำให้เกิดการเลื่อนไปทางซ้ายจะใช้การพิจารณาชุดของตัวอักษรที่ตรงกันของ text และ pattern ที่อยู่ทางขวาของตำแหน่งที่เกิดการ Miss-match เรียกว่า Suffix และชุดของตำแหน่งตั้งแต่เกิดการ Miss-match ไปทางซ้ายเรียกว่า Prefix ซึ่งเรียกว่า Good suffix shift

สำหรับกรณี Bad character shift ยกตัวอย่างดังภาพที่ 1 เมื่อมี text คือ abbababacba และ pattern คือ babac เมื่อทำการเปรียบเทียบจะพบว่าเกิดการ Miss-match กันที่ ตำแหน่งที่ 4 ของ text คือตัวอักษร b ใน text เพราะฉะนั้นตำแหน่งตัวอักษร b ที่ปรากฏอยู่ใน pattern คือตำแหน่ง 0 และ 2 พิจารณาจากค่าที่มากที่สุดคือ 2 เพราะฉะนั้นจะทำการเลื่อนไปทางขวา 2 ตำแหน่ง

ตำแหน่ง	0	1	2	3	4	5	6	7	8	9	...
text	a	b	b	a	b	a	b	a	c	b	a
pattern	b	a	b	a	c						
			b	a	b	a	c				

ภาพที่ 1: ตัวอย่าง Bad character shift

บางครั้งหากพิจารณาตามกรณี Bad character shift จะทำให้เกิดการเลื่อนไปทางซ้ายมีดังภาพที่ 2 พิจารณาตำแหน่งการ Miss-match คือตำแหน่งที่ 2 ตัวอักษรที่ Miss-match คือ a ตำแหน่งทางขวาที่ดีที่สุดของ a ใน pattern คือ 1 และ 3 จะทำให้เกิดการเลื่อนไปทางซ้ายซึ่งจะไม่สามารถทำได้ ในกรณีนี้ต้องใช้การพิจารณา Good suffix shift

ตำแหน่ง	0	1	2	3	4	5	6	7	8	9	...
text	a	b	a	a	b	a	b	a	c	b	a
pattern	c	a	b	a	b						
			c	a	b	a	b				

ภาพที่ 2: ตัวอย่าง Bad character shift เลื่อนไปทางซ้าย

สำหรับการพิจารณา Good suffix shift จากตัวอย่างภาพที่ 3 พิจารณา suffix ในกรณีนี้คือ a_b ซึ่งเมื่อนำ suffix ไปเทียบกับชุดตัวอักษรใน prefix พบว่าไม่ปรากฏอยู่ใน prefix จะทำการเลื่อนไปทางขวาเท่ากับจำนวนของตัวอักษรของ pattern คือ 5

ตำแหน่ง	0	1	2	3	4	5	6	7	8	9	...
text	a	b	c	a	b	a	b	a	c	b	a
pattern	c	b	a	a	b						
						c	b	a	a	b	

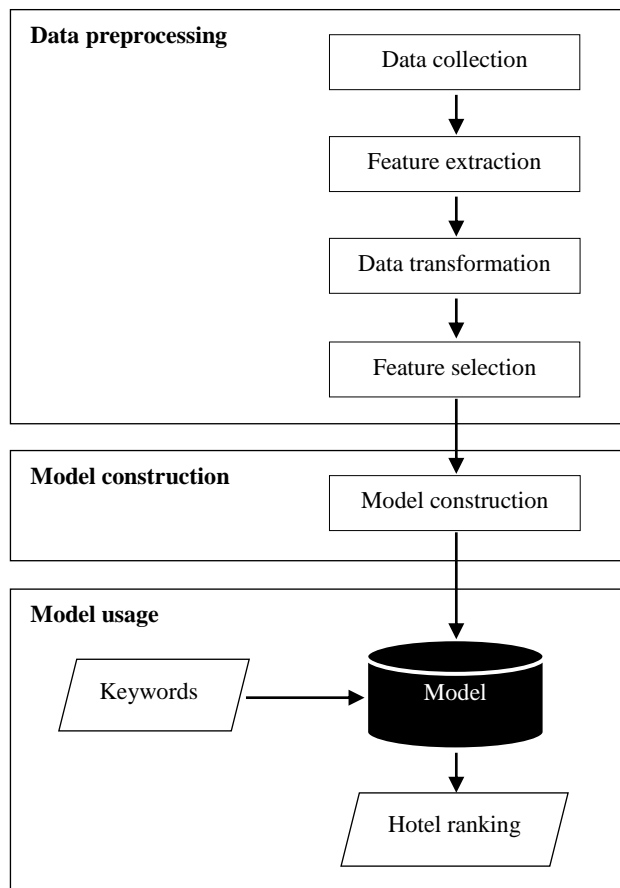
ภาพที่ 3: ตัวอย่าง Good suffix shift

3. ขั้นตอนการดำเนินงานวิจัย

งานวิจัยนี้ได้แบ่งการทำงานทั้งหมดออกเป็น 3 ส่วนดังภาพที่ 4 ได้แก่ 1.ขั้นตอนการเตรียมข้อมูล (Data preprocessing) เป็นการเตรียมข้อมูลให้เหมาะกับโมเดลที่ใช้ในการเรียนรู้ 2.ขั้นตอนการสร้างโมเดล (Model construction) เป็นขั้นตอนการสร้างโมเดลเพื่อนำไปใช้ในการค้นหาโรงแรม และ 3. ขั้นตอนการนำโมเดลไปใช้ (Model usage) ขั้นตอนนี้เป็นขั้นตอนที่นำคำค้นหาจากผู้ใช้งานมาประมวลผลเพื่อแสดงรายชื่อโรงแรมที่มีความสอดคล้องกับความต้องการของผู้ใช้มากที่สุดตามลำดับจากมากไปหาน้อยจากโมเดลที่สร้างขึ้นมาในขั้นตอนก่อนหน้า

ในส่วนของขั้นตอนการเตรียมข้อมูลประกอบด้วย การเก็บข้อมูล (Data collection) การสกัดคุณลักษณะ (Feature extraction) การแปลงข้อมูล (Data transformation) และการคัดเลือกคุณลักษณะ (Feature selection) ในการเก็บข้อมูลตัวอย่าง ผู้วิจัยได้ใช้ข้อมูลการแสดงความคิดเห็นจากเว็บไซต์ www.tripadvisor.com โดยเลือกโรงแรมตัวอย่างในจังหวัดภูเก็ตจำนวน 15 โรงแรม และเลือกข้อมูลความคิดเห็นทั้งหมด 293 ความคิดเห็น คัดเลือกความคิดเห็นที่มีคะแนนความคิดเห็นตั้งแต่ 3 คะแนนขึ้นไป โดยในแต่ละความคิดเห็นประกอบด้วย

คะแนนด้านต่างๆของโรงแรมซึ่งแบ่งเป็น 5 กลุ่มคือ คุณภาพการนอน (Sleep quality) คุณภาพสถานที่ (Location) คุณภาพห้องพัก (Rooms) คุณภาพการบริการ (Services) และคุณภาพความสะอาด (Cleanliness)



ภาพที่ 4: ขั้นตอนการดำเนินงานวิจัย

ขั้นตอนถัดมาคือการสกัดคุณลักษณะซึ่งเป็นการสร้างคำศัพท์จากความคิดเห็นทั้งหมด โดยเก็บคำศัพท์ที่ปรากฏในความคิดเห็นและทำให้อยู่ในรูปของรากศัพท์โดยตัด -s, -es, -ly, -ing, -ed และตัดคำบุพบท คำสรรพนามต่างๆ เช่น in, on, the, a, is, am, are, I, you, we, they เป็นต้น แล้วจึงนำมาสร้างเป็นพจนานุกรมซึ่งประกอบด้วยคำศัพท์ทั้งหมด 3,071 คำ หลังจากนั้นทำการแปลงข้อมูลให้อยู่ในรูปแบบ 0,1 ดังตารางที่ 2 โดย 0 หมายถึง คำศัพท์นั้นไม่ปรากฏในความคิดเห็น และ 1 หมายถึง คำศัพท์นั้นปรากฏในความคิดเห็น เนื่องจากคุณลักษณะหรือคำศัพท์ที่สกัดได้มีจำนวนมากซึ่งอาจทำให้มีผลต่อประสิทธิภาพของโมเดลที่สร้าง จึงได้มีการใช้เทคนิคการ

คัดเลือกคุณลักษณะที่สำคัญเพื่อลดจำนวนคุณลักษณะหรือคำศัพท์ลง ซึ่งงานวิจัยนี้เลือกใช้วิธีการคัดเลือกคุณลักษณะด้วยเทคนิค Correlation feature selection (CFS) ทำให้ได้คุณลักษณะหรือคำศัพท์ที่สำคัญทั้งหมด 85 คำ

ตารางที่ 2: ตัวอย่างการแปลงข้อมูล

ความคิดเห็นที่	คำศัพท์					
	abdominal	ability	able	about	...	zone
1	0	1	1	1	...	0
2	1	1	0	0	...	1
3	0	0	0	1	...	0
.
.
.
293	1	0	1	0	...	0

ขั้นตอนต่อไปคือการสร้างโมเดลเพื่อนำไปใช้ในการค้นหาโรงแรมจากคำค้นหา ซึ่งงานวิจัยนี้เลือกใช้เทคนิคตัวจำแนกแบบเบย์อย่างง่าย เพื่อหาความน่าจะเป็นของคำศัพท์แต่ละคำในการปรากฏอยู่ในความคิดเห็นของแต่ละโรงแรม P(Vocab|Hotel) แสดงดังตารางที่ 3

ตารางที่ 3: ค่าความน่าจะเป็นของคำศัพท์ที่ปรากฏอยู่ในความคิดเห็นแต่ละโรงแรม

โรงแรม	คำศัพท์				
	access	always	apartment	...	work
1	0.2516	0.2516	0.3080	...	0.2422
2	0.2914	0.2731	0.2464	...	0.3087
3	0.2363	0.2548	0.2275	...	0.2548
.
.
.
15	0.2231	0.2037	0.1746	...	0.2134

ขั้นตอนสุดท้ายคือการนำโมเดลที่ได้มาใช้ในการค้นหาโรงแรมโดยผู้ใช้ป้อนคำค้นหา (Keywords) หลังจากนั้นโมเดลจะทำการหาค่าความน่าจะเป็นของแต่ละโรงแรมจากคำที่ใช้ใน

การค้นหาและแสดงผลเป็นลำดับของโรงแรมทั้ง 15 โรงแรมตามค่าความน่าจะเป็นจากมากไปน้อย

การหาค่าความน่าจะเป็นของโรงแรมเพื่อนำมาจัดลำดับการแสดงผลนั้น คำนวณตามสมการที่ 6

$$P(\text{Hotel}|\text{Keywords})=P(\text{Hotel}) \prod_i P(\text{Keywords}_i|\text{Hotel}) \quad (6)$$

เมื่อ $P(\text{Hotel}|\text{Keywords})$ คือ ความน่าจะเป็นของโรงแรมเมื่อมี keyword ต่างๆปรากฏอยู่ในความคิดเห็นของโรงแรมนั้น

$P(\text{Hotel})$ คือ ความน่าจะเป็นของความคิดเห็นที่มีต่อโรงแรมนั้น

$P(\text{Keywords}_i|\text{Hotel})$ คือ ความน่าจะเป็นของคำค้นหาเมื่อคำค้น หานั้น ปรากฏ ใน โรงแรม ซึ่ง หา ได้ จาก ค่า $P(\text{Vocab}|\text{Hotel})$ ตามตารางที่ 3

4. การทดลองและผลการทดลอง

ในการทดลองนี้ได้ทำการเปรียบเทียบการค้นหาโรงแรมด้วยตัวจำแนกแบบเบย์อย่างง่ายกับวิธีการค้นหาแบบตรรกะและวิธีการค้นหาแบบบอยเซอร์-มัวร์ โดยทำการทดสอบประสิทธิภาพการทำงานของโมเดล 2 ด้าน คือ การเปรียบเทียบผลการค้นหาโรงแรมและเวลาที่ใช้ในการค้นหาโรงแรม

สำหรับคำค้นหาที่ใช้ในการทดลองประกอบด้วย 7 ชุดคำค้นหา จำแนกตามคุณภาพด้านต่างๆ คือ คุณภาพการนอน คุณภาพสถานที่ คุณภาพห้องพัก คุณภาพการบริการ คุณภาพความสะอาดและอีก 2 ชุดคำค้นหาคือผสมคำทุกกลุ่มไว้ด้วยกัน (ทุกด้าน(1)และทุกด้าน(2)) ดังตารางที่ 4

ตารางที่ 4: คำที่ใช้ค้นหาในแต่ละด้านจำแนกตามหมวดต่างๆ

หมวดการค้นหา	คำค้นหา
การนอน	sleep private
สถานที่	beautiful beach bay club restaurant
ห้องพัก	clean balcony kitchen
การบริการ	massage spa shuttle tv
ความสะอาด	wonderful clean
ทุกด้าน (1)	private beach balcony shuttle clean
ทุกด้าน (2)	clean beach sea private balcony

ในส่วนของการวัดประสิทธิภาพการค้นหาในงานวิจัยนี้ได้กำหนดค่า Rank_accuracy ซึ่งเป็นเปอร์เซ็นต์ของค่าความ

ถูกต้องในการเรียงลำดับโรงแรมจากการค้นหาสามารถคำนวณได้ตามสมการที่ 7 โดยที่ค่าน้ำหนัก w_i ของลำดับโรงแรมที่อยู่ลำดับแรกของผลการค้นหาจะมีค่าสูงที่สุด และจะมีค่าลดลงเรื่อยๆตามลำดับ

$$\text{Rank Accuracy}(\%) = \left(\frac{\sum_{i=1}^{\text{rank}} \text{Rating}_{\text{avg}} \times w_i}{\sum_{i=1}^{\text{rank}} \text{Rating}_{\text{max}} \times w_i} \right) \times 100 \quad (7)$$

เมื่อ $\text{Rating}_{\text{avg}}$ คือคะแนนเฉลี่ยจากความคิดเห็นของแต่ละโรงแรม

$\text{Rating}_{\text{max}}$ คือคะแนนเต็มจากความคิดเห็นของแต่ละโรงแรม โดยมีคะแนนเต็มที่ 5 คะแนน

w_i คือค่าน้ำหนักของแต่ละลำดับ

rank คือจำนวนลำดับทั้งหมดจากการค้นหาโรงแรมมีค่าเท่ากับ 15

จากตารางที่ 5 พบว่าค่า Rank_accuracy ในการค้นหาด้วยตัวจำแนกแบบเบย์อย่างง่ายมีค่าสูงกว่าวิธีการค้นหาแบบตรรกะและแบบบอยเซอร์-มัวร์ ยกเว้นชุดคำค้นหาทุกด้าน(2) ตัวจำแนกแบบเบย์อย่างง่ายมีค่า Rank_accuracy เท่ากับ 81.17% วิธีการค้นหาแบบบอยเซอร์-มัวร์มีค่า Rank_accuracy เท่ากับ 82.24% แต่เมื่อวิเคราะห์ผลการทดลองจากค่าเฉลี่ยของ Rank_accuracy ตัวจำแนกแบบเบย์อย่างง่ายมีค่า Rank_accuracy สูงกว่าวิธีการค้นหาแบบตรรกะและวิธีการค้นหาแบบบอยเซอร์-มัวร์ โดยที่ตัวจำแนกแบบเบย์อย่างง่ายมีค่าเฉลี่ยของ Rank_accuracy เท่ากับ 80.93% ซึ่งสูงกว่าวิธีการค้นหาทั้งสองวิธีอยู่ที่ 72.41% และ 79.08% ตามลำดับ นอกจากนี้พบว่าตัวจำแนกแบบเบย์อย่างง่ายให้ผลการจัดเรียงลำดับในด้านการนอนได้ดีกว่าวิธีอื่นอย่างเห็นได้ชัด

ตารางที่ 5: ค่า Rank_accuracy ในแต่ละวิธี

	ตัวจำแนกแบบเบย์	การค้นหาแบบ	การค้นหาแบบ
	อย่างง่าย	ตรรกะ	บอยเซอร์-มัวร์
ทุกด้าน (1)	81.07	71.11	80.81
ทุกด้าน (2)	81.17	72.38	82.24
การนอน	77.99	64.84	74.11
สถานที่	79.23	71.84	79.15
ห้องพัก	79.98	74.08	79.15
บริการ	81.24	70.64	80.97
ความสะอาด	86.21	80.65	85.63
ค่าเฉลี่ย	80.93	72.41	79.80

ในส่วนการวัดประสิทธิภาพทางด้านเวลาในการค้นหาโรงแรม ผลที่ได้ปรากฏว่าตัวจำแนกแบบเบย์อย่างง่ายใช้เวลาเฉลี่ยในการประมวลผลน้อยที่สุดคือ 0.00768 วินาที วิธีการค้นหาแบบตรรกะใช้เวลา 0.00815 วินาที และวิธีการค้นหาแบบบอยเซอร์-มัวร์ใช้เวลา 1.38037 วินาที แสดงได้ตามตารางที่ 6

ตารางที่ 6: เวลาที่ใช้ในการค้นหาโรงแรมตามวิธีค้นหา

	เวลาในการค้นหา (วินาที)		
	ตัวจำแนกแบบเบย์ อย่างง่าย	การค้นหาแบบ ตรรกะ	การค้นหาแบบ บอยเซอร์-มัวร์
ทุกด้าน (1)	0.01882	0.01125	1.76695
ทุกด้าน (2)	0.00590	0.00781	1.89873
การนอน	0.00544	0.00582	0.73581
สถานที่	0.00565	0.00956	1.87108
ห้องพัก	0.00604	0.00682	1.12114
บริการ	0.00612	0.00924	1.58301
ความสะอาด	0.00578	0.00657	0.68584
ค่าเฉลี่ย	0.00768	0.00815	1.38037

ในการทดลองนี้เมื่อวิเคราะห์ผลการทดลองจากค่าเฉลี่ยของ Rank_accuracy พบว่าตัวจำแนกแบบเบย์อย่างง่ายให้ผลลัพธ์ใกล้เคียงกับวิธีการค้นหาแบบบอยเซอร์-มัวร์ แต่เมื่อดูจากความเร็วในการประมวลผลแล้ววิธีการค้นหาแบบบอยเซอร์-มัวร์ใช้เวลาในการประมวลผลมากกว่าตัวจำแนกแบบเบย์อย่างง่าย เมื่อวิเคราะห์ผลการทดลองด้านเวลาในการประมวลผลตัวจำแนกแบบเบย์อย่างง่ายและการค้นหาแบบตรรกะจะใช้เวลาในการประมวลผลใกล้เคียงกัน แต่เมื่อพิจารณาค่าเฉลี่ยของ Rank_accuracy ตัวจำแนกแบบเบย์อย่างง่ายมีค่าที่สูงกว่าอย่างเห็นได้ชัด

5. บทสรุป

งานวิจัยนี้ได้ประยุกต์ใช้เทคนิคการทำเหมืองข้อมูลความคิดเห็นโดยใช้ตัวจำแนกแบบเบย์อย่างง่ายเพื่อทำการจัดลำดับการค้นหาของโรงแรม โดยทำการทดลองเปรียบเทียบกับวิธีการค้นหาแบบตรรกะและวิธีการค้นหาแบบบอยเซอร์-มัวร์ จากการทดลองพบว่าวิธีการค้นหาโดยใช้ตัวจำแนกแบบเบย์อย่างง่ายนั้นให้ผลที่ดีกว่าทั้งในด้านความถูกต้องในการจัดลำดับโรงแรมที่ต้องการค้นหา และประสิทธิภาพด้านเวลา

ในการประมวลผล เมื่อพิจารณาเปรียบเทียบค่า Rank_accuracy ที่ได้จากการทดลอง พบว่าตัวจำแนกแบบเบย์อย่างง่ายมีความถูกต้องอย่างมากเมื่อเปรียบเทียบกับวิธีการค้นหาแบบตรรกะและมีค่าใกล้เคียงกับวิธีการค้นหาแบบบอยเซอร์-มัวร์ และเมื่อเปรียบเทียบประสิทธิภาพเวลาในการประมวลผลของตัวจำแนกแบบเบย์อย่างง่ายกับวิธีการค้นหาแบบบอยเซอร์-มัวร์ พบว่าตัวจำแนกแบบเบย์อย่างง่ายใช้เวลาในการประมวลผลเร็วกว่ามาก

ผลการทดลองในบทความนี้ถือว่าเป็นที่น่าพอใจ แต่ยังคงต้องพัฒนาวิธีการคัดเลือกคุณลักษณะเพื่อให้ได้คำศัพท์ที่ได้มีคุณภาพมากยิ่งขึ้น ซึ่งจะทำให้ผลการจัดอันดับดียิ่งขึ้นอีกด้วย

เอกสารอ้างอิง

- [1] จุติมา เกษมศรีธนาวัฒน์ และ ธนสนี เพียรตระกูล, “การจำแนกความคิดเห็นโดยใช้ตัวจำแนกแบบเบย์ร่วมกับการเลือกคุณลักษณะด้วยอัลกอริทึมรีลีฟ,” *National conference on Computer Information Technology 2011 (CIT2011) and UniNet Network Operation and Management Workshop 2011 (UniNOMS2011)*, 26-28 มกราคม 2554.
- [2] C. Troussas, M. Virvou, K. J. Espinosa, K. Llaguno, and J. Caro, “Sentiment analysis of Facebook statuses using Naive Bayes classifier for language learning,” *Fourth International Conference on Information, Intelligence, Systems and Applications (IISA)*, July 10-12, 2013.
- [3] B. Liu, E. Blasch, Y. Chen, D. Shen, and G. Chen, “Scalable sentiment classification for Big Data analysis using Naïve Bayes Classifier,” *IEEE International Conference on Big Data*, Oct 6-9, 2013.
- [4] W. Fan, S. Sun, and G. Song, “Sentiment Classification for Chinese Netnews Comments Based on Multiple Classifiers Integration,” *Fourth International Joint Conference on Computational Sciences and Optimization (CSO)*, Apr 15-19, 2011.
- [5] Tom Michael, *Machine learning*, McGraw-Hill: USA, 1997.
- [6] ศุภชัย ตั้งวงศ์สานต์, *ระบบการจัดเก็บและการสืบค้นสารสนเทศด้วยคอมพิวเตอร์*, พิมพ์ครั้งที่ 1, โรงพิมพ์พิทักษ์การพิมพ์: กรุงเทพฯ, 2551.

การแยกกลุ่มคำถามเพื่อค้นหาคำตอบโดยใช้เว็บเชิงความหมายและการทำเหมืองข้อมูล

Question Classification for Answer Searching Using Semantic Web and Data Mining

สุพจน์ บัวเลิง (Supote Bualerng)¹และ วรรัตน์ สงฆ์เป็น (Wararat Songpan)²

¹คณะกรรมการบัญชีและการจัดการ มหาวิทยาลัยมหาสารคาม

²ภาควิชาวิทยาการคอมพิวเตอร์คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น

¹supote.b@acc.msu.ac.th, ²wararat@kku.ac.th

บทคัดย่อ

งานวิจัยนี้มีจุดประสงค์เพื่อพัฒนาหาต้นแบบการค้นหาคำตอบที่เกี่ยวข้องกับคำถามแบบอัตโนมัติ ซึ่งกระบวนการทำงานของระบบได้นำคำหลักที่เป็นตัวแทนของประโยคของคำถามและคำตอบมาทำการวัดความคล้ายคลึงของคำเชิงความหมาย (Semantic Similarity) โดยอาศัยฐานข้อมูลคำศัพท์เชิงความหมายในเวิร์ดเน็ต (Wordnet) จากนั้นได้ทำการแยกกลุ่มของคำหลักโดยใช้หลักการของเลซิกอลเชน (Lexical Chain) เพื่อไปหาทฤษฎีความสัมพันธ์เชื่อมโยงคำหลักของคำถามและคำตอบโดยใช้อัลกอริทึมแอฟพริออริ (Apriori algorithm) ในการประมวลผล ผลลัพธ์คำตอบที่ได้จากการค้นหาคำตอบที่เชื่อมโยงกับคำถามภาษาไทยแบบเชิงความหมายจากคำหลักของคำถามและคำตอบที่ได้ทำการทดสอบทั้งหมดจำนวน 200 เรคคอร์ด ได้วัดประสิทธิภาพโดยรวมของความถูกต้อง (F-Measure) โดยเฉลี่ยของกลุ่มตัวอย่างเท่ากับ 84.36% ซึ่งระบบสามารถทำงานโดยการค้นหาคำถามและสามารถหาประโยคคำตอบที่เกี่ยวข้องได้ซึ่งตรงกับความต้องการ

คำสำคัญ: ความคล้ายคลึงของคำเชิงความหมาย เวิร์ดเน็ต เลซิกอลเชน แอฟพริออริ ประสิทธิภาพการค้นหา

Abstract

The objective of this research is to develop a prototype of question and answer system automatically. The process of system starts from searched by keywords of question and answer sentence and compared semantic similarity between the keywords based on wordnet vocabulary. The vocabulary are classify into categorizing keywords by Lexical Chain method and link the

relationship between questions and answer results that are analyzed by Association Rules with Apriori algorithm. The experimental results found that the system can search meaningful answers relevant with the question sentences from 200 records. The average of F-Measure that is overall of Precision and recall as 84.36%. The overall of system performance can question search and give the answers correctly.

Keywords: semantic similarity, wordnet, lexical chain, Apriori, F-measure.

1. บทนำ

ปัจจุบันนี้การเกิดสารสนเทศไม่ว่าจะเป็นคำถามคำตอบที่มากมายและยังไม่ได้มีการรวบรวมเพื่อแยกแยะประเภทเพื่อให้ง่ายต่อการนำไปใช้งาน จะเห็นได้ว่าเมื่อมีผู้ถามคำถามบนเว็บ จะมีคำถามอื่นๆที่มีความใกล้เคียงและเป็นการคำถามเดียวกัน ทำให้ผู้ดูแลเว็บต้องมีการตอบคำถามซ้ำๆ จึงเกิดความซ้ำซ้อนและเสียเวลาสำหรับการตอบคำถามดังกล่าว ดังนั้นเว็บเชิงความหมาย (Semantic Web) [1,2] ซึ่งเป็นเทคโนโลยีที่ใช้ในการจัดการจัดเก็บและนำเสนอเนื้อหาแบบมีโครงสร้างรวมถึงสามารถที่จะวิเคราะห์ จำแนกหรือจัดแบ่งได้ว่าคำที่ปรากฏในเอกสารมีความสัมพันธ์กับเอกสารอื่นๆ กล่าวคือเป็นการจัดเก็บและนำเสนอแบบมีลำดับชั้น (Hierarchy) โดยมีการแยกกลุ่มปัญหาโดยการเลือกทำเหมืองข้อมูล (Data Mining) จากเนื้อหาของเว็บจะจำกัดกลุ่มตามหัวข้อ (Content) ของคำถามปัญหาต่างๆ ซึ่งการทำงานแบบเว็บเชิงความหมายจะช่วยให้คำถามหรือกลุ่มปัญหาต่างๆ ที่ใกล้เคียงหรือมีลักษณะกลุ่มเดียวกันได้รวมที่จะจัดหมวดหมู่ได้ง่ายต่อการค้นหาและช่วยอำนวยความสะดวก

สะดวกต่อผู้ใช้งานและผู้สนใจในการใช้งานสารสนเทศดังกล่าว

ดังนั้นงานวิจัยจึงนำเสนอสถาปัตยกรรมและพัฒนาเป็นต้นแบบของระบบการทำงานการแยกกลุ่มคำถามเพื่อการค้นหาคำตอบโดยอาศัยเว็บเชิงความหมายโดยมุ่งเน้นการสร้างการแยกกลุ่มคำถามบนเว็บและการทำเหมืองข้อมูล

2. วรรณกรรมและงานวิจัยที่เกี่ยวข้อง

2.1 เว็บเชิงความหมาย (Semantic Web)

เว็บเชิงความหมาย (Semantic Web) [1,2] เป็นการจัดการข้อมูลในลักษณะของการเชื่อมโยงความสัมพันธ์ของข้อมูลในระดับเมตาดาต้า (Metadata) เพื่อทำการอ่านข้อมูลแบบออนไลน์ทำให้เครื่องคอมพิวเตอร์สามารถเข้าใจ ความหมายของข้อมูลต่างๆ ได้ว่าเป็นอะไรมาจากข้อมูลส่วนไหนของชุดข้อมูล ทำให้คอมพิวเตอร์ สามารถนำข้อมูลที่ได้นั้นไปประมวลผลได้โดยอัตโนมัติ ยกตัวอย่างเช่น ถ้าคอมพิวเตอร์พบข้อมูลว่า X เป็นน้องของ Y และ Z เป็นลูกของ Y คอมพิวเตอร์จะสามารถรับรู้ได้เองว่า Z เป็นหลานของ X เป็นต้น

2.2 การตัดคำ (Word Segmentation)

การตัดคำ (Word segmentation) [3,4,5] คือการแบ่งตัวอักษรจากข้อความ(String)เพื่อหาขอบเขตของแต่ละหน่วยคำ (Morpheme) เนื่องจากส่วนใหญ่ภาษาไทยมีการเขียนในลักษณะที่ติดกันโดยไม่มีการใช้เครื่องหมายวรรคตอนกั้นระหว่างคำเหมือนภาษาอังกฤษซึ่งใช้ช่องว่าง (Space) กั้นระหว่างคำ แต่ภาษาไทยจะมีการเว้นวรรคเป็นระยะดังนั้นการตัดคำในงานวิจัยฉบับนี้ได้ใช้หลักการตัดคำโดยใช้พจนานุกรมโดยวิธีการเทียบคำที่ยาวที่สุด (Longest Matching) ซึ่งเป็นเทคนิคอย่างหนึ่งที่มีการตัดคำด้วยวิธีทาง Heuristic อย่างหนึ่ง และเมื่อทดสอบกับระบบผู้วิจัยพบว่าให้ประสิทธิภาพที่ดีเมื่อใช้กับภาษาไทย

2.3 การสรุปใจความสำคัญ (Text Summarization)

การสรุปใจความสำคัญ (Text Summarization) [6] เป็นเทคนิคที่ใช้ในการสรุปเนื้อหาของเอกสารโดยมีขั้นตอนในการสรุปเนื้อหาของเอกสารอยู่ 3 ขั้นตอนคือ ขั้นตอนการวิเคราะห์คำ ขั้นตอนการหาประเด็นสำคัญ และขั้นตอนการสังเคราะห์ประโยค โดยสามารถแบ่งวิธีการสรุปใจความสำคัญตามวิธีการ

สรุปแบบภาพรวม (Abstraction) ซึ่งเป็นการวิเคราะห์คำจากต้นฉบับด้วยหลักการทางภาษาศาสตร์ (Semantic) โดยเป็นการถอดความหมายหรือแปลความหมายจากต้นฉบับผลที่ได้มีประสิทธิภาพและความถูกต้องมากกว่าการใช้วิธี Extraction ซึ่งวิธีการ Abstraction ประกอบด้วย 3 ขั้นตอนดังนี้

1) การสกัดคำหลัก

เป็นการหาคำหลักของเอกสารโดยอาศัยจากฐานข้อมูลคำศัพท์ WordNet ในการคัดเลือกคำหลักจากเอกสาร การคัดเลือกคำหลักจะคัดเลือกเอาเฉพาะคำที่สามารถสื่อความหมายได้เท่านั้น

2) การจัดหมวดหมู่ของคำ (Lexical Chain)

เป็นการนำคำหลักที่สกัดได้มาทำการจัดกลุ่มใหม่หรือทำการสร้าง Chain ของคำศัพท์โดยอาศัยความสัมพันธ์ของคำเชิงความหมาย จากฐานข้อมูลคำศัพท์เว็บริดเน็ตที่ได้นิยามความสัมพันธ์ของคำศัพท์ไว้มาทำการจัดหมวดหมู่ของคำหลักซึ่งมีหลักการทำงานดังภาพที่ 1 ดังต่อไปนี้

```

REPEAT
  READ next word

  IF word is suitable for lexical analysis (see section 3.2.1) THEN
    CHECK for chains within a suitable span
    (up to 3 intermediary sentences, and no limitation on returns):
    CHECK thesaurus for relationships (section 3.2.2).
    CHECK other knowledge sources
    (situational, general words, proper names).
  IF chain relationship is found THEN
    INCLUDE word in chain.
    CALCULATE chain so far
    (allow one transitive link).
  END IF
  IF there are words that have not formed a chain for a suitable
  number of sentences (up to 3) THEN
    ELIMINATE words from the span.
  END IF
  CHECK new word for relevance to existing chains that
  are suitable for checking.
  ELIMINATE chains that are not suitable for checking.
  END IF
END REPEAT
    
```

ภาพที่ 1: ลำดับขั้นตอนการสร้าง Lexical Chain

3) การคำนวณค่าน้ำหนัก Lexical Chain

เป็นการคำนวณและให้ค่าน้ำหนักในแต่ละ Chain ว่าสมาชิกใน Chain มีความสัมพันธ์กันมากน้อยเพียงใด การหาค่าน้ำหนักของ Lexical Chain สามารถคำนวณได้จากสมการ

$$S_k = \left(\sum_{m=1}^{P_k+1} tf_{mk} R_{mk} W_{mk} \right) H_k \tag{1}$$

โดย

S_k คือ คำนับหน้าของ Lexical Chain ที่ k

m คือ คำหลักที่ m

f_{mk} คือ ความถี่ของการเกิดคำหลักที่ m ที่เกิดขึ้นซ้ำใน

Lexical Chain ที่ k

R_{mk} คือ น้ำหนักของความสัมพันธ์ของคำที่ m ใน

Lexical Chain ที่ k

W_{mk} คือ น้ำหนักของคำหลักที่ m ที่เกิดขึ้นซ้ำใน Lexical

Chain ที่ k

H_k คือ ค่า Homogeneity Index ซึ่งคำนวณได้จาก

$$H_k = 1 - \frac{P_k}{L_k} \quad (2)$$

โดย

L_k คือ จำนวนสมาชิกที่เกิดแตกต่างกันใน Lexical

Chain ที่ k

P_k คือ จำนวนคำหลักที่เป็นสมาชิกทั้งหมดใน Lexical Chain

ที่ k

2.4 กฎความสัมพันธ์ (Association Rules)

การหาความสัมพันธ์ [7] คือ การได้มาของกฎความสัมพันธ์โดยการหารูปแบบที่เกิดขึ้นบ่อยคู่กัน และเรียกว่า frequent pattern และความสัมพันธ์ที่เกิดขึ้น เรียกว่า association ของกลุ่มไอเท็มจากข้อมูลที่อยู่ในรูปแบบ transaction ผลลัพธ์ที่ได้จะอยู่ในรูปแบบของกฎความสัมพันธ์ (Association Rules)

$$\text{item1} \rightarrow \text{item2}$$

ในงานวิจัยนี้ได้ใช้อัลกอริทึมเพื่อหาความสัมพันธ์ คือ แอปพริออริ (Apriori) หากต้องการทราบความสัมพันธ์ของกฎจะต้องคำนวณหาค่าสนับสนุน ค่าความเชื่อมั่นและการหาค่าสหสัมพันธ์ระหว่างข้อมูล เพื่อทราบความสัมพันธ์ของข้อมูลทั้งสองว่ามีความสัมพันธ์มากน้อยเพียงใด

การคำนวณหาค่าสนับสนุน Support

$$\text{Support}(X \rightarrow Y) = P(XUY) \quad (3)$$

คือ ค่าความน่าจะเป็นของจำนวนของค่าข้อมูล X เกิดขึ้นคู่กับค่าข้อมูล Y ซึ่งมีค่าอยู่ระหว่าง [0-1] ถ้าค่าใกล้ 1 หมายถึงมีความสัมพันธ์กันมาก

การคำนวณหาค่าความเชื่อมั่น Confidence (Conf)

$$\text{Confidence}(X \rightarrow Y) = P(XUY)/P(X) \quad (4)$$

คือ ค่าความน่าจะเป็นของจำนวนของค่าข้อมูล X เกิดขึ้นคู่กับค่าข้อมูล Y ตามมาเสมอโดยที่มีลำดับเหตุการณ์ของข้อมูลเข้ามาเกี่ยวข้องซึ่งมีค่าอยู่ระหว่าง [0-1] ถ้าค่าใกล้ 1 หมายถึงมีความสัมพันธ์กันตามลำดับของข้อมูลมาก

การคำนวณหาค่าสหสัมพันธ์ (Lift)

$$\text{Lift}(X \rightarrow Y) = \frac{P(XUY)}{P(X)P(Y)} \quad (5)$$

หมายถึง ค่าวัดความสัมพันธ์ระหว่างค่า X และ Y ว่ามีความสัมพันธ์กันจริงหรือไม่ ถ้าค่า Lift มีค่าเท่ากับ 0 และน้อยกว่าหรือเท่ากับ 1 แสดงว่าข้อมูล X และ Y ไม่สัมพันธ์กัน หากมีค่ามากกว่า 1 แสดงว่าข้อมูล X และ Y สัมพันธ์กันจริง

2.5 การวัดประสิทธิภาพการสืบค้นของระบบ

การวัดประสิทธิภาพการสืบค้น [8] เป็นการตรวจสอบความถูกต้องของการสืบค้นเอกสาร โดยพิจารณาจากระดับความแม่นยำ (Precision) และความครบถ้วน (Recall) รวมไปถึงประสิทธิภาพโดยรวมของการสืบค้น (F-Measure) ตามสมการต่อไปนี้

$$F - \text{measure} = \frac{2RP}{R+P} \quad (6)$$

$$P = \frac{A}{A+B} \times 100 \quad (7)$$

$$R = \frac{A}{A+C} \times 100 \quad (8)$$

โดย

P คือ ค่าความถูกต้อง (Precision)

R คือ ค่าความครบถ้วน (Recall)

A คือ จำนวนเอกสารที่ถูกต้องและถูกสืบค้น

B คือ จำนวนเอกสารที่ไม่ถูกต้องและถูกสืบค้น

C คือ จำนวนเอกสารที่ถูกต้องแต่ไม่ถูกสืบค้น

2.6 งานวิจัยที่เกี่ยวข้อง

งานวิจัย [1,2,8] มุ่งเน้นเพื่อการออกแบบแนวคิดวิธีการสรุปลักษณะของเอกสารที่มีประสิทธิภาพและถูกต้อง โดยมีการใช้เทคโนโลยีทางด้านเชิงความหมาย (Semantic) เข้ามาช่วยในการคัดเลือกสกัดคำหลัก โดยอาศัยฐานข้อมูลคำศัพท์เวิร์คเดตในการพิจารณาร่วมด้วย แล้วทำการจัดหมวดหมู่ของคำโดยใช้เทคนิคการสร้าง Lexical Chain มาวิเคราะห์หาความสัมพันธ์ของคำหลักที่สกัดจากเอกสาร และคัดเลือกผลลัพธ์จากลำดับ

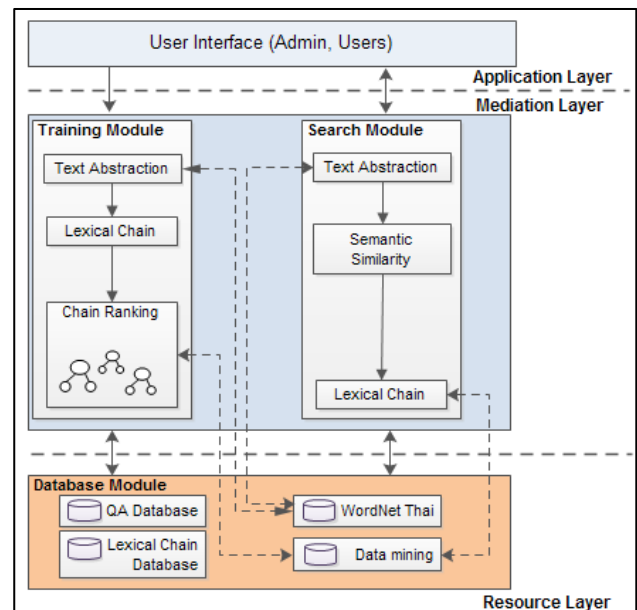
ความสำคัญของกลุ่มคำโดยใช้ค่าถ่วงน้ำหนักเชิงความหมายประกอบ โดยใช้ออนโทโลยี เพื่อลดเวลาในการพิจารณาคัดเลือกข้อมูลข่าวสารอิเล็กทรอนิกส์ โดยอาศัยเทคโนโลยีทางด้านเชิงความหมาย (Semantic) งานวิจัย [3,4,5] ได้นำเสนอวิธีการตัดคำภาษาไทยซึ่งเป็นขั้นตอนพื้นฐานที่สำคัญสำหรับระบบงานทางด้านการประมวลผลภาษาธรรมชาติ (Natural Language Processing) โดยการตัดคำภาษาไทยประกอบด้วยปัญหาความกำกวมและปัญหาคำที่ไม่ปรากฏในพจนานุกรมเนื่องมาจากลักษณะรูปแบบประโยคในภาษาไทยจะเขียนเรียงติดต่อกันไปโดยไม่มีเว้นวรรคระหว่างคำทำให้ระบบไม่ทราบขอบเขตที่ชัดเจนของคำจึงส่งผลกระทบต่อระบบงานที่เกี่ยวข้องกับการประมวลผลภาษาธรรมชาติ เช่น ระบบสืบค้นข้อมูล ดังนั้นจึงได้มุ่งเน้นการศึกษาและพัฒนาวิธีการค้นหาคำที่ไม่ปรากฏในพจนานุกรมโดยได้นำเสนอการตัดคำไทยโดยการเทียบสายอักษรโดยการค้นหาจากพจนานุกรมซึ่งได้ประยุกต์การสร้างกราฟโดยวิธีการหาเส้นทางที่สั้นที่สุดเพื่อหาเซตของกลุ่มคำที่ตัดได้จากข้อความนำเข้าและประยุกต์ใช้การระบุชนิดของคำตามหลักภาษาไทยเพื่อสร้างกฎสำหรับการหาขอบเขตของคำที่ทับกันและขอบเขตคำที่ไม่ปรากฏในพจนานุกรมงานวิจัย [9] ได้นำเสนอการปรับปรุงประสิทธิภาพการทำเหมืองเว็บโดยใช้เทคนิคการจัดหมวดหมู่และกฎความสัมพันธ์โดยใช้เว็บไซต์มหาวิทยาลัยขอนแก่นเป็นกรณีศึกษาโดยใช้แคชเซิร์ฟเวอร์ (Cache Server) ตามสถาปัตยกรรมแม่แบบ SP-Cache ในการหาความสัมพันธ์ระหว่างกลุ่มสามารถทำได้โดยเลือกข้อมูล URL ที่ต้องการหาความสัมพันธ์โดยหาความสัมพันธ์ (Support) ที่เป็นค่าบอกความเป็นไปได้ของข้อมูลในการเกิดร่วมกันของกฎค่าความเชื่อมั่น (Confidence) เป็นค่าที่บอกถึงสัดส่วนระหว่างค่าสนับสนุนในการเกิดร่วมกันของกฎและค่า Lift ซึ่งเป็นค่าที่บอกถึงสัดส่วนระหว่างค่าสนับสนุนและค่าความเชื่อมั่นในการร่วมกันของกฎ และงานวิจัย [10] งานมุ่งเน้นไปที่การเพิ่มประสิทธิภาพในการสืบค้นรูปภาพบนอินเทอร์เน็ตเพื่อให้ได้ผลการสืบค้นที่ตรงกับความต้องการของผู้ใช้ โดยการนำเอาแนวความคิดของเว็บเชิงความหมาย (Semantic Web) เพื่อช่วยในค้นหารูปภาพจากแนวความคิดดังกล่าวผลลัพธ์จากการค้นหาที่มีความแม่นยำสูงและมีประสิทธิภาพมากยิ่งขึ้น

3. วิธีการดำเนินการวิจัย

วิธีการดำเนินการศึกษาวิจัยได้ทำการสกัดคำจากคำถามและคำตอบที่เก็บในระบบฐานข้อมูลที่เป็นภาษาไทยจากข้อมูลเดิมที่มีอยู่แล้วและมีการจัดเก็บในระบบฐานข้อมูล MySQL และทำการสุ่มมาเพื่อทำการวัดประสิทธิภาพของระบบจำนวน 200 เรคคอร์ด โดยเลือกเอาประโยคคำถามในเว็บบอร์ดที่มีคนอ่านมากที่สุดมาดำเนินเป็นกลุ่มตัวอย่าง

3.1 การออกแบบสถาปัตยกรรมของระบบ

ระบบได้ทำการออกแบบแบ่งเป็น 3 ส่วนคือ



ภาพที่ 2: สถาปัตยกรรมระบบการค้นหา

1) Application Layer คือ ส่วนที่ทำหน้าที่ติดต่อกับผู้ใช้งานระบบ จะรับข้อมูลจากผู้ใช้เพื่อนำข้อมูลไปประมวลผลในระบบและแสดงผลลัพธ์กลับไปยังผู้ใช้งาน

2) Mediation Layer คือส่วนการทำงานประมวลผลของระบบ จะรับข้อมูลที่ผู้ใช้งานส่งเข้ามาทำการประมวลผลตามขั้นตอนวิธีต่างๆ และส่งผลลัพธ์กลับไปยังผู้ใช้งานระบบและจัดเก็บในฐานข้อมูล ใน Mediation Layer นี้จะแบ่งการทำงานออกเป็น 2 โมดูลหลัก ได้แก่

- Training Module จะทำหน้าที่ในการเรียนรู้คำของกลุ่มประโยคคำถาม และคำของกลุ่มประโยคคำตอบจากข้อมูลชุดตัวอย่างให้กับระบบ โดยเริ่มจากการรับคำมาจากผู้ใช้งานเข้ามาสู่ขั้นตอน Text Abstraction โดยการตัดคำแล้วนำคำที่ตัดได้ไปค้นหาความหมายในฐานข้อมูลคำศัพท์เวิร์ดเน็ต (WordNet) เพื่อ

ทำการคัดเอาเฉพาะคำที่เป็นคำนาม และคำกริยา ที่สื่อความหมายเท่านั้น แล้วทำการจัดกลุ่มคำที่มีความหมายใกล้เคียงกันให้อยู่ในกลุ่มเดียวกันในขั้นตอน การทำ Lexical Chain และทำการคำนวณค่าน้ำหนักของ Lexical Chain ว่ากลุ่มคำใดที่มีค่าน้ำหนักมากที่สุดจะถูกจัดไว้ลำดับต้นๆ ในขั้นตอน Chain Ranking และนำกลุ่มคำของคำถามและกลุ่มคำของคำตอบไปหาความสัมพันธ์หรือรูปแบบการเกิดของคำด้วยโปรแกรม weka และเมื่อได้รูปแบบความสัมพันธ์ระหว่างคำถามกับคำตอบที่เกิดขึ้น แล้วทำการบันทึกกฎความสัมพันธ์ที่ได้ลงในระบบฐานข้อมูล ในส่วนของ Resource Layer

- Search Module จะทำหน้าที่ในการค้นคำตอบให้กับผู้ใช้งาน จากข้อมูลที่ได้รับเข้ามาจากผู้ใช้งาน ในส่วนนี้จะทำงานเหมือนกับ Training Module ทุกอย่าง จะต่างกันตรงที่ Search Module จะนำกลุ่มคำที่ได้จากขั้นตอนการทำ Lexical Chain ไปเทียบกับ กฎความสัมพันธ์ในระบบฐานข้อมูลที่ได้จากขั้นตอนการทำงานของ Training Module เพื่อจะทำการดึงเอาคำตอบกลับไปแสดงผลให้กับ ผู้ใช้งาน ในส่วนของ User Interface

3) Resource Layer คือส่วนที่จัดเก็บเก็บข้อมูลที่ได้จากการประมวลผลและข้อมูลเอกสารที่ใช้ในการประมวลผลซึ่งในส่วนนี้จะประกอบไปด้วย 4 ระบบฐานข้อมูลดังนี้

- ระบบฐานข้อมูล WordNet สำหรับภาษาไทยใช้เก็บคำศัพท์เชิงความหมายเพื่อใช้ในขั้นตอน Text Abstraction ของทั้ง Training Module และ Search Module รวมทั้งในส่วนของ การ จัดกลุ่มคำ Lexical Chain ด้วย

- ระบบฐานข้อมูล Lexical Chain จะเก็บกลุ่มคำที่มีความหมายใกล้เคียงกันไว้ด้วยกันรวมทั้งเก็บค่าน้ำหนักของกลุ่มคำในส่วนของการทำ Chain Ranking ไว้จาก ทั้ง Training Module และ Search Module เช่นกัน

- ระบบฐานข้อมูล Data Mining ทำหน้าที่เก็บกฎความสัมพันธ์ที่ได้จากการหาค่าความสัมพันธ์ระหว่างคำถามกับคำตอบ โดยใช้โปรแกรม Weka [11] เพื่อนำไปใช้ในส่วนของ Search Module การหาคำของคำถามตั้งต้น โดยการเลือกมาจากคำที่เกิดความถี่มากที่สุดคือ 16 จากการทำงานของ Training Module ในส่วนของ Chain Ranking จนถึงน้อยที่สุดคือ 4 จะได้ทั้งหมดจำนวน 15 คำ แล้วกำหนดให้เป็นตัวแปร Q1 จนถึง Q15 การหาคำของคำตอบตั้งต้นโดยการเลือกมาจากคำตอบที่

ตอบคำถามทั้ง 15 ข้อที่เกิดจำนวนของคำตอบมากที่สุดในแต่ละคำถามและตัดกลุ่มคำที่ซ้ำกันออก จะได้ทั้งหมดจำนวน 29 คำ ที่เป็นกลุ่มคำตอบที่มากที่สุดในแต่ละคำถาม แล้วกำหนดให้เป็นตัวแปร A1 จนถึง A25 ตามลำดับเพื่อนำไปประมวลผลในโปรแกรม weka เพื่อหาคำตอบที่ต้องการ ดังภาพที่ 3

Best rules found:	
1.Q8=4543 1 ==>A17=95650 1 conf:(1) <lift:(14.75)> lev:(0.01) [0] conv:(0.93)	
2.A17=95650 8 ==>Q8=4543 1 conf:(0.13) <lift:(14.75)> lev:(0.01) [0] conv:(0.99)	
3.Q9=4 4 ==>A19=39455 1 conf:(0.25) <lift:(9.83)> lev:(0.01) [0] conv:(0.97)	
4.A19=39455 3 ==>Q9=4 1 conf:(0.33) <lift:(9.83)> lev:(0.01) [0] conv:(0.97)	
5.Q10=81559 5 ==>A22=34997 1 conf:(0.2) <lift:(7.87)> lev:(0.01) [0] conv:(0.97)	
6.A22=34997 3 ==>Q10=81559 1 conf:(0.33) <lift:(7.87)> lev:(0.01) [0] conv:(0.96)	
7.Q14=95650 5 ==>A22=34997 1 conf:(0.2) <lift:(7.87)> lev:(0.01) [0] conv:(0.97)	
8.A22=34997 3 ==>Q14=95650 1 conf:(0.33) <lift:(7.97)> lev:(0.01) [0] conv:(0.96)	
9.Q3=85108 1 ==>A4=90293 1 conf:(1) <lift:(7.38)> lev:(0.01) [0] conv:(0.86)	
10.A4=90293= 16 ==>Q3=85108 1 conf:(0.06) <lift:(7.38)> lev:(0.01) [0] conv:(0.99)	

ภาพที่ 3: ผลลัพธ์จากโปรแกรม Weka ที่ใช้ในการวิเคราะห์

- ระบบฐานข้อมูล QA จะเก็บคำถาม คำตอบที่เป็นกลุ่มตัวอย่างในการดำเนินการวิจัยในครั้งนี้ จะถูกนำไปใช้ใน ส่วนของ Training Module เพื่อทำการสอนให้ระบบเรียนรู้คำในกลุ่มตัวอย่าง และนำคำตอบไป แสดงในส่วนของ User Interface โดยผ่าน Search Module ในการดำเนินการ

3.2 ขั้นตอนการพัฒนาฐานความรู้เชิงความหมาย

ขั้นตอนการดำเนินการออกเป็น 4 ขั้นตอน ดังนี้

1) การสกัดความรู้ การตัดคำและการสกัดคำหลัก

คำถามคำตอบตัวอย่างภาษาไทยที่ใช้เป็นตัวอย่างการทดลองครั้งนี้จะถูกรวบรวมและจัดเก็บไว้ในฐานข้อมูลเพื่อให้ง่ายต่อการนำมาสกัดคำหลักโดยระบบฐานข้อมูลที่ใช้คือ ฐานข้อมูล MariaDB จะถูกตัดข้อความและสกัดคำหลักออกมาโดยใช้โปรแกรมตัดคำสำหรับข้อความภาษาไทย Swath คือโปรแกรมแบ่งคำภาษาไทย ซึ่งพัฒนาโดยใช้เทคนิคการเรียนรู้ด้วยเครื่องคอมพิวเตอร์ (Machine Learning)

2) การจัดกลุ่มความรู้โดยการสร้าง Lexical Chain

Lexical Chain เป็นการนำคำหลักที่สกัดได้มาทำการจัดกลุ่มใหม่ หรือทำการสร้าง Chain ของคำศัพท์โดยอาศัยความสัมพันธ์ของคำเชิงความหมาย โดยอ้างอิงฐานข้อมูลคำศัพท์เวิร์ดเน็ตในการหาความสัมพันธ์ของคำซึ่งได้ปรับปรุงการสร้าง Lexical Chain

3) การคำนวณค่าน้ำหนักของLexical Chain

เป็นการคำนวณและให้ค่าน้ำหนักในแต่ละ Chain ว่าสมาชิกใน Chain มีความสัมพันธ์กันมากน้อยเพียงใด ทั้งนี้ในการหาค่าน้ำหนักของความสัมพันธ์จะแบ่งออกเป็น ระดับ 3 ดังนี้

ความสัมพันธ์แบบ Extra-strong เป็นความสัมพันธ์แบบเหนียวแน่นมาก จะเกิดขึ้นเมื่อเกิดคำเดิมซ้ำบ่อยๆ

ความสัมพันธ์แบบ Strongเป็นความสัมพันธ์แบบเหนียวแน่น จะเกิดขึ้นเมื่อคำมีความหมายเฉพาะเจาะจง,คำพ้องความหมาย และคำที่มีความหมายครอบคลุม

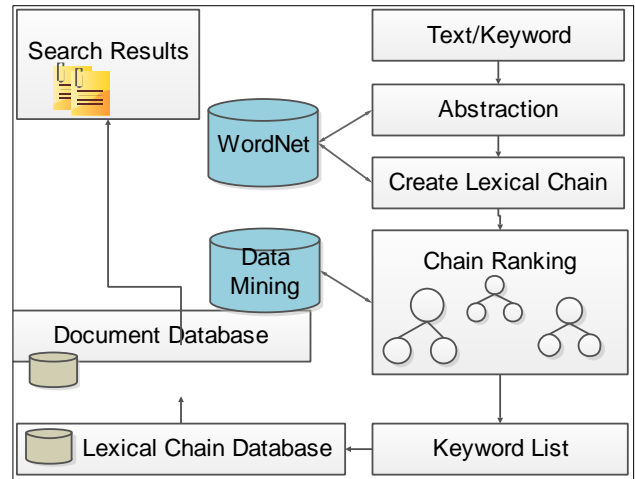
ความสัมพันธ์แบบ Medium-strongเป็นความสัมพันธ์แบบปานกลาง จะเกิดขึ้นเมื่อเส้นเชื่อมระหว่าง Synset อยู่ระหว่างชั้นที่ 3 ถึง ชั้นที่ 2

ตารางที่ 1:ความสัมพันธ์และค่าน้ำหนัก

ประเภทความสัมพันธ์		ค่าน้ำหนัก
Extra-strong	Repetition	1.0
Strong	Synonym	0.9
	Hypernym, Hyponym, Meronym, Holonym, troponym (Level 1)	0.6
Medium-strong	Hypernym, Hyponym, Meronym, Holonym, troponym (Level 2 - 3)	0.2
Medium-strong	Hypernym, Hyponym, Meronym, Holonym, troponym (Level 3 - 5)	0.1

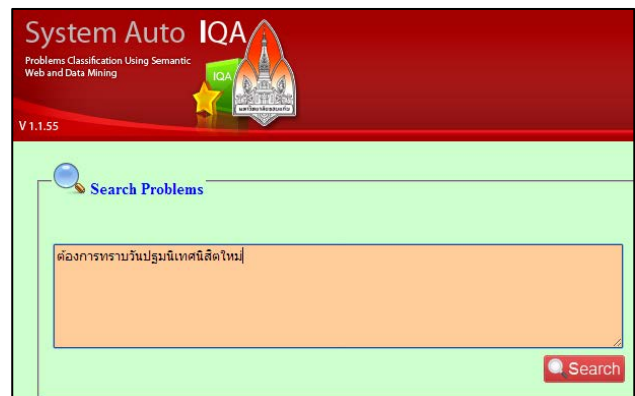
3.3 ขั้นตอนการพัฒนาาระบบสืบค้นการจัดกลุ่มคำถามแบบเว็บเชิงความหมาย

การพัฒนาาระบบสืบค้นการจัดกลุ่มคำถามแบบเว็บเชิงความหมายจะใช้ข้อมูลที่ได้ทำการหาตัวแทนกลุ่มคำถามคำตอบและทำการจัดกลุ่มของ Lexical Chain เรียบร้อยแล้วที่จัดเก็บไว้ในฐานข้อมูลมาทำการสืบค้นข้อมูลเชิงความหมายโดยอาศัยฐานข้อมูลคำศัพท์ WordNet ในการหาความหมายของคำค้น กระบวนการพัฒนาระบบการสืบค้นนั้นมีขั้นตอนและกระบวนการของระบบ ดังภาพประกอบที่ 4



ภาพที่ 4: ขั้นตอนการทำงานของระบบค้นหา

จากภาพที่ 4 การทำงานของระบบสืบค้นเริ่มจากการหาความหมายของคำค้นในขั้นตอนนี้จะนำคำหรือประโยคที่ทำการสืบค้นมาทำการหาความหมายของคำและสรุปใจความสำคัญของประโยคโดยอาศัยฐานข้อมูลคำศัพท์ WordNet ในการหาความหมายของคำและสรุปใจความสำคัญของประโยคโดยใช้หลักการ Lexical Chain ในการสรุปใจความสำคัญของประโยคและสืบค้นข้อมูลจากฐานข้อมูลการทำเหมืองข้อมูลที่ประมวลผลไว้เรียบร้อยแล้วทำการแสดงผลการค้นหาดังภาพที่ 5



ภาพที่ 5: ตัวอย่างการพัฒนาาระบบการค้นหาคำตอบอัตโนมัติ

ตารางที่ 2:ตัวอย่างคำถามเพื่อทดสอบการค้นหา ตามกฎข้อที่ 1

คำค้นหา
ต้องการทราบวันปฐมนิเทศนิสิตใหม่

ตารางที่ 3: ตัวอย่างคำตอบจากระบบการค้นหา

คำตอบที่#1 1586
วันที่ 13 พ.ค. 54 ปฐมนิเทศปริญญาโทและปริญญาเอก
วันที่ 25 พ.ค. 54 ระดับปริญญาตรีเทียบเข้า พบอาจารย์ที่ปรึกษาและดำเนินการทำเรื่องเทียบ โอนรายวิชา
วันที่ 26 พ.ค.54 ระดับปริญญาตรีเทียบเข้า ปฐมนิเทศ

<p>คำตอบที่#2 1754</p> <p>การเทียบสามารถเทียบโอนได้แต่ต้องดูรายละเอียด ขอให้คิดต่อที่เจ้าหน้าที่คณะกรรมการ บัญชีและการจัดการครับ</p>
<p>คำตอบที่#3 1594</p> <p>วันที่ 25 พ.ค. 54 ระดับปริญญาตรีเทียบเข้า พบอาจารย์ที่ปรึกษาและดำเนินการทำเรื่องเทียบโอนรายวิชา</p> <p>วันที่ 26 พ.ค.54 ระดับปริญญาตรีเทียบเข้า ปริญญาโทศ ม.เก่า</p>
<p>คำตอบที่#4 1592</p> <p>วันที่ 25 พ.ค. 54 ระดับปริญญาตรีเทียบเข้า พบอาจารย์ที่ปรึกษาและดำเนินการทำเรื่องเทียบโอนรายวิชา อาคาร 3 ม.เก่าค่ะ</p> <p>วันที่ 26 พ.ค.54 ระดับปริญญาตรีเทียบเข้า ปริญญาโทศ กำลังจะปิดรับสมัคร บนเว็บแล้ว</p>

4. ผลการดำเนินงานวิจัย

จากการวิเคราะห์โดยใช้วิธีการทำเหมืองความสัมพันธ์ มีกลุ่มคำถามและกลุ่มคำตอบที่นำมาวิเคราะห์ในการประมวลผล ได้กลุ่มออกมาทั้งหมด 10 กฎ และมีกฎความสัมพันธ์ในการสืบค้นดังนี้

4.1 กฎที่สามารถนำมาเพื่อใช้พัฒนาระบบสืบค้นหาคำตอบ

จากการวิเคราะห์เราจะสนใจเฉพาะกฎที่ขึ้นต้นด้วย Q ที่อยู่ฝั่งซ้ายของลูกศรจากกฎที่ได้จะเห็นได้ว่ามี 5 กฎที่น่าสนใจและจะสามารถนำไปทดสอบการใช้งานได้ ดังต่อไปนี้

ตารางที่ 4:กลุ่มคำ Q8(ปริญญานิเทศ) →A17(เทียบ)

กฎ	ค่าConf	ค่าLift
1. Q8=4543 1 ==> A17=95650 1	1	14.75

ตัวอย่างกฎที่1 หมายถึง ถ้าสืบค้นจากกลุ่มคำถาม “ปริญญานิเทศ” แล้วจะได้กลุ่มคำตอบ “เทียบ”มีค่าความเชื่อมั่นเท่ากับ 100% ที่จะเกิดตามกฎและกฎมีความสัมพันธ์กันจริงมากกว่า 1 คือ 14.75

ตารางที่ 5:กลุ่มคำ Q9(ของ) →A19(เรื่อง)

กฎ	ค่าConf	ค่าLift
3. Q9=4 4 ==> A19=39455 1	0.25	9.83

ตัวอย่างกฎที่3 หมายถึง ถ้าสืบค้นจากกลุ่มคำถาม “ของ” แล้วจะได้กลุ่มคำตอบ “เรื่อง” มีค่าความเชื่อมั่นเท่ากับ 25% ที่จะเกิดตามกฎและกฎมีความสัมพันธ์กันจริงมากกว่า1 คือ 9.83

ตารางที่ 6:กลุ่มคำ Q10(ปี) →A22(ปฏิทิน)

กฎ	ค่าConf	ค่าLift
5. Q10=81559 5 ==> A22=34997 1	0.20	7.87

ตัวอย่างกฎที่5 หมายถึง ถ้าสืบค้นจากกลุ่มคำถาม “ปี” แล้วจะได้กลุ่มคำตอบ “ปฏิทิน” มีค่าความเชื่อมั่นเท่ากับ 20% ที่จะเกิดตามกฎและกฎมีความสัมพันธ์กันจริงมากกว่า1 คือ 7.87

ตารางที่ 7:กลุ่มคำ Q14(เทียบ) →A22(ปฏิทิน)

กฎ	ค่าConf	ค่าLift
7. Q14=95650 5 ==> A22=34997 1	0.20	7.87

ตัวอย่างกฎที่7 หมายถึง ถ้าสืบค้นจากกลุ่มคำถาม “เทียบ” แล้วจะได้กลุ่มคำตอบ “ปฏิทิน” มีค่าความเชื่อมั่นเท่ากับ 20% ที่จะเกิดตามกฎและกฎมีความสัมพันธ์กันจริงมากกว่า1 คือ 7.87

ตารางที่ 8:กลุ่มคำ Q3(รับ) →A4(ได้)

กฎ	ค่าConf	ค่าLift
9. Q3=85108 1 ==> A4=90293 1	1	7.38

ตัวอย่างกฎที่ 9 หมายถึง ถ้าสืบค้นจากกลุ่มคำถาม “รับ” แล้วจะได้กลุ่มคำตอบ “ได้” มีค่าความเชื่อมั่นเท่ากับ 100% ที่จะเกิดตามกฎและกฎมีความสัมพันธ์กันจริงมากกว่า 1 คือ 7.38

จากกฎความสัมพันธ์ที่สามารถนำไปใช้งานได้จริง จำนวน 5 กฎ และทำให้ทราบรูปแบบการหาคำตอบ ของคำถามตามข้อมูลที่ได้จากการวิจัย โดยทดสอบแบ่งออกเป็น 2 รูปแบบ โดยแบบที่ 1 เป็นกลุ่มคำถามแบบหลายประโยชน์ขยารวมกัน ได้ทำการทดสอบ 4 กลุ่มคำถาม ดังตารางที่ 7 ต่อไปนี้

ตารางที่ 9: วัดประสิทธิภาพกลุ่มคำถามรูปแบบที่ 1 แบบหลายประโยชน์ขยารวมกัน

กลุ่มคำถามเกี่ยวกับ	Precision	Recall	F-Measure
ระยะเวลาการรับสมัคร	78.54%	100%	86.92%
วันเวลาในการ	89.83%	100%	94.05%

ปฐมนิเทศ			
การกู้ยืม กยศ.	68.50%	100%	80.11%
การเทียบโอน รายวิชา	73.76%	100%	83.92%
ค่าเฉลี่ย	77.66%	100%	86.25%

และรูปแบบที่ 2 จากกลุ่มคำถามแบบคำถามเดียว หรือ ประโยคเดียว โดยมีการวัดประสิทธิภาพการสืบค้นด้วยค่าความถูกต้องในการสืบค้น (Precision) โดยเฉลี่ย 73.96% การวัดค่าความครบถ้วนในการสืบค้นด้วยค่าความครบถ้วน (Recall) โดยเฉลี่ย 100% และการวัดค่าประสิทธิภาพการสืบค้นด้วยค่า F-measure มีค่าโดยเฉลี่ย 82.47% ดังนั้นเมื่อวัดประสิทธิภาพค่าเฉลี่ยโดยรวมของระบบโดยวัดเป็นค่าเฉลี่ยทั้งสองรูปแบบจะได้ค่า F-measure เท่ากับ 84.36%

5. สรุป

การวิจัยนี้ได้ใช้คำถามและคำตอบภาษาไทยจำนวน 200 เรคคอร์ดที่มีจำนวนการอ่านมากที่สุดไปจนถึงน้อยที่สุดมาใช้ในการทดลองจะประกอบด้วยขั้นตอนหลัก 2 ขั้นตอน ขั้นตอนที่ 1 ขั้นตอนสรุปใจความสำคัญของคำถามเพื่อหาตัวแทนของประโยคโดยใช้หลักการ Lexical Chain ขั้นตอนที่ 2 ขั้นตอนการสืบค้นข้อมูลเชิงความหมาย การวิจัยนี้ได้พัฒนาและออกแบบระบบสืบค้นคำถามภาษาไทยเชิงความหมายที่ใช้ข้อมูลจากการประมวลผลในขั้นตอนสรุปใจความสำคัญของประโยคคำถามซึ่งจะเก็บค่าน้ำหนักของแต่ละ Chain ที่เป็นตัวแทนของประโยคมาใช้ในการสืบค้นผลของคำตอบที่ได้ถือว่าค่อนข้างมีประสิทธิภาพดี ด้วยการวัดประสิทธิภาพการสืบค้นด้วยค่า F-measure เฉลี่ยของกลุ่มตัวอย่างของประโยคคำถามทั้งหมด 2 รูปแบบ 5 กลุ่มมีค่าเฉลี่ยโดยรวมเท่ากับ 84.36% รวมทั้งระบบสามารถทำงานและประมวลผลนำคำตอบที่ได้มาตอบรวมทั้งยังมีคำตอบอื่นๆที่อยู่ในกลุ่มเดียวกันมาตอบด้วยตรงกับวัตถุประสงค์และสมมุติฐานที่ตั้งไว้ การหาคำตอบถือว่าค่อนข้างมีประสิทธิภาพดี เพราะระบบสามารถทำงานและประมวลผลนำคำตอบที่เกี่ยวข้องกันมาตอบรวมในคำถามที่มีความหมายเดียวกันได้ถูกต้องและมีประสิทธิภาพ ทำให้ลดปัญหาการตอบคำถามที่มีความซ้ำซ้อนเป็นจำนวนมาก

เอกสารอ้างอิง

- [1] อรรถพรณ อุไรเรืองพันธ์. การสรุปเอกสารเชิงความหมายโดยใช้ออนโทโลยี. [รายงานการศึกษาระดับปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ] ขอนแก่น.: บัณฑิตวิทยาลัย มหาวิทยาลัยขอนแก่น; 2552.
- [2] วิชิตา โชติรัตน์, ศุภศิ บุญรอด และ ศจีมาจ ณ วิเชียร, “การพัฒนาฐานความรู้ออนโทโลยี สำหรับวิเคราะห์ข่าวออนไลน์โดยอัตโนมัติ” *วารสารเทคโนโลยีสารสนเทศพระจอมเกล้าพระนครเหนือ* ปีที่ 7 ฉบับที่ 14 กรกฎาคม-ธันวาคม 2554 หน้า 13-18.
- [3] ฐาปนี เสงสนั่นกุล .การตัดคำภาษาไทย .[วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิตสาขาวิชาวิทยาการคอมพิวเตอร์]. ขอนแก่น:บัณฑิตวิทยาลัย มหาวิทยาลัยขอนแก่น; 2549.
- [4] สายัณห์ เทพแดง. การปรับปรุงประสิทธิภาพของการตัดคำภาษาไทยด้วยเทคนิคการจดจำนิพจน์ระนาบ. [วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิตสาขาวิชาวิทยาการคอมพิวเตอร์]. บัณฑิตวิทยาลัย มหาวิทยาลัยธรรมศาสตร์; 2553.
- [5] สุรศักดิ์ ตั้งสกุล, ฐาปนี เสงสนั่นกุล, “การตัดคำภาษาไทยสำหรับข้อความในพีพธิภกันซ์ปลาน้ำจืด” *วารสารวิทยาศาสตร์บูรพา* ปีที่ 16 ฉบับที่ 1 2554 หน้า 84-93.
- [6] กิตยศิริ ช่อเจียงนวลวรรณ สุนทรภิชัย .ขั้นตอนวิธีสรุปข้อคิดเห็นภาษาไทยโดยใช้ออนโทโลยี. ใน: National Conference on Computer Information Technologies 2010 (CIT2010); 2553 มกราคม 16-13; มหาวิทยาลัยบูรพา วิทยาเขตสารสนเทศจังหวัดจันทบุรี จันทบุรีกรุงเทพฯ.: มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ; .2554 หน้าที่ .18-13
- [7] วิถีพิเศษไทย. การทำเหมืองข้อมูล. [ออนไลน์] 2554 [อ้างเมื่อ 14 สิงหาคม 2554]. จาก <http://th.wikipedia.org/wiki>
- [8] GuiZhen Yang. “On Precision and recall of multi-attribute data extraction from semistructured sources” *Proceeding of the IEEE International Conference on Data Mining(ICDM)*, Nov. 19-22,2003, pp.395-402.
- [9] สมฤทัย ปั้นพานิช. การทำเหมืองเว็บโดยเทคนิคการจัดหมวดหมู่และกฎความสัมพันธ์. [วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ]. ขอนแก่น:บัณฑิตวิทยาลัย มหาวิทยาลัยขอนแก่น; 2550.
- [10] ธนกร หวังพัฒน์วงศ์, อนันท์ ไกรเสวกวิสัย และ สราวุธ ราษฎร์นิยม. ระบบค้นหารูปภาพโดยใช้หลักการเว็บเชิงความหมาย The Image Search Engine using Semantic Web. กรุงเทพฯ: สำนักงานคณะกรรมการการอุดมศึกษา; 2553.
- [11] WEKA. [ออนไลน์] 2554 [อ้างเมื่อ 14 สิงหาคม 2554]. จาก <http://www.cs.waikato.ac.nz/ml/weka/>

ระบบจัดหมวดหมู่สถานที่ท่องเที่ยวด้วยคำอธิบายภาษาไทยแบบอัตโนมัติผ่านเว็บเซอร์วิส

Automatic Tourist Attraction Categorization System using Thai Description via Web Service

ชูพันธุ์ รัตน์โกคา (Choopan Rattanapoka)¹ และเมธาวี สุทธิกุล (Mathawee Sutikun)²

¹ภาควิชาเทคโนโลยีวิศวกรรมอิเล็กทรอนิกส์ วิทยาลัยเทคโนโลยีอุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

²ภาควิชาเทคโนโลยีวิศวกรรมอิเล็กทรอนิกส์ วิทยาลัยเทคโนโลยีอุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

choopanr@kmutnb.ac.th, mathawee.1990@hotmail.com

บทคัดย่อ

บทความนี้เสนอการออกแบบและพัฒนาระบบจัดหมวดหมู่สถานที่ท่องเที่ยวแบบอัตโนมัติ จากคำอธิบายสถานที่ท่องเที่ยวที่เป็นภาษาไทยด้วยภาษาจาวา โดยข้อมูลคำอธิบายสถานที่ท่องเที่ยว จะถูกส่งเข้ามายังระบบผ่านช่องทางบริการของเว็บเซอร์วิส ซึ่งหมวดหมู่สถานที่ท่องเที่ยวที่ระบบรองรับในการจัดหมวดหมู่ มีทั้งหมด 4 หมวดหมู่ คือ วัด ทะเล ภูเขา และที่พัก ด้วยการประยุกต์ใช้วิธีการเรียนรู้ของเครื่องผ่านโปรแกรมเวก้า จากผลการทดลองโดยการทดสอบกับสถานที่ท่องเที่ยวในจังหวัดภูเก็ตพบว่า ระบบมีความแม่นยำในการจัดหมวดหมู่สถานที่ท่องเที่ยวอยู่ที่ 95%

คำสำคัญ: เวก้า การจำแนกหมวดหมู่ การเรียนรู้ของเครื่อง เว็บเซอร์วิส

Abstract

This paper presents a design and implementation of tourism spots categorized system using Thai description of the tourist attractions. The information of tourist attraction input to our system via web service. Then, our system automatically classifies them into the properly category by using machine learning algorithm via Weka. Our system supports 4 tourism spots categories: temple, beach, mountain and accommodation. In the experiments, we use tourist attractions in Phuket and found out that our system has about 95% accuracy for the tourism attraction categorization.

Keyword: Weka, Classification, Machine Learning, Web Service

1. บทนำ

ปัจจุบันอุตสาหกรรมการท่องเที่ยวในประเทศไทย ถือว่าเป็นรายได้หลักอย่างหนึ่งของประเทศ การรวบรวมข้อมูลและการประชาสัมพันธ์สถานที่ท่องเที่ยวให้นักท่องเที่ยวสามารถค้นหาข้อมูลได้อย่างสะดวกนั้นควรจะมีการจัดหมวดหมู่ให้กับสถานที่ท่องเที่ยวอย่างเป็นระบบ และการที่จะให้ได้มาซึ่งสถานที่ท่องเที่ยวและข้อมูลเกี่ยวกับสถานที่ต่างๆ เข้ามาเก็บในฐานข้อมูลนั้น จะสะดวกมากขึ้นถ้าเปิดโอกาสให้บุคคลทั่วไปสามารถช่วยกันเพิ่มข้อมูลต่างๆ เกี่ยวกับสถานที่ท่องเที่ยวได้ แต่อย่างไรก็ตามเมื่อมีการนำเข้าสู่ข้อมูลเกี่ยวกับสถานที่ท่องเที่ยวเป็นจำนวนมาก ซึ่งผู้นำเข้าข้อมูลสถานที่ท่องเที่ยวอาจจะไม่ได้จำแนกหมวดหมู่สถานที่ท่องเที่ยวมาด้วย หรือใส่หมวดหมู่ผิด จะทำให้เกิดภาระกับผู้ดูแลระบบที่หน้าทีตรวจสอบและจัดหมวดหมู่สถานที่ท่องเที่ยว ยิ่งไปกว่านั้นถ้าผู้ดูแลระบบไม่เคยไปสถานที่นั้นๆ มาก่อน จะยิ่งทำให้การตัดสินใจในการตรวจสอบและจำแนกหมวดหมู่ของสถานที่ท่องเที่ยวลำบากมากยิ่งขึ้น

ดังนั้นเพื่อให้การจัดสถานที่ท่องเที่ยวไปตามหมวดหมู่ที่ถูกต้อง อีกทั้งยังลดภาระและข้อผิดพลาดให้กับผู้ดูแลระบบในการจัดหมวดหมู่สถานที่ท่องเที่ยว งานวิจัยนี้จึงต้องการนำเสนอการออกแบบและพัฒนาระบบจัดหมวดหมู่สถานที่ท่องเที่ยวแบบอัตโนมัติด้วยวิธีการเรียนรู้ของเครื่องผ่านโปรแกรม Weka จากคำอธิบายสถานที่ท่องเที่ยวภาษาไทย โดยให้บริการในรูปแบบของเว็บเซอร์วิส โดยขอบเขตของระบบต้นแบบใช้สำหรับสถานที่ท่องเที่ยวภายในจังหวัดภูเก็ต และรองรับหมวดหมู่สถานที่ท่องเที่ยวได้ทั้งหมด 4 หมวดหมู่ คือ วัด ทะเล ภูเขา และที่พัก

2. วรรณกรรมที่เกี่ยวข้อง

การจัดแบ่งหมวดหมู่ข้อความ [1], [2] คือ การระบุประเภทของเอกสารหรือข้อความแบบอัตโนมัติโดยอิงจากตัวเนื้อหาในเอกสารหรือข้อความนั้น ๆ การนำการจัดหมวดหมู่ข้อความมาช่วยใช้การจัดหมวดหมู่ของสถานที่ท่องเที่ยว จะเป็นการค้นหารูปแบบและความสัมพันธ์ในชุดข้อมูล โดยมีการนำเทคนิคในการเรียนรู้ของเครื่อง (Machine Learning) มาประยุกต์ใช้และสร้างแบบจำลองเพื่อทำนายข้อมูลของสถานที่ท่องเที่ยวที่เข้ามาใหม่ ว่าควรจัดอยู่ในหมวดหมู่ใด

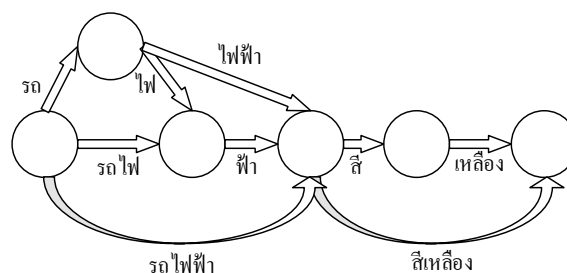
งานวิจัยที่เกี่ยวข้องกับการจัดหมวดหมู่สถานที่ท่องเที่ยวที่พบมีมากมายเช่น [3] ที่นำข้อมูลเมตาดาต้า (Meta data) ของรูปภาพสถานที่ท่องเที่ยวซึ่งอยู่ในรูปแบบข้อความ ที่ได้จากเว็บไซต์ผู้ให้บริการอัปโหลดรูปภาพฟลิคเกอร์ (Flickr) มาทำการจำแนกและระบุสถานที่ เพื่อนำมาวิเคราะห์ และแนะนำตารางการเดินทางภายในหนึ่งวันว่าไปเที่ยวสถานที่ใดได้บ้างหรือใน [4] เป็นงานวิจัยที่มุ่งเน้นการจัดทำพจนานุกรมทางด้านภูมิศาสตร์ ซึ่งมีการจัดหมวดหมู่สถานที่ทางด้านภูมิศาสตร์ เช่น ภูเขา แม่น้ำ โบราณสถาน ฯลฯ โดยนำข้อมูลจากฟลิคเกอร์ และวิกิพีเดีย มาใช้เป็นข้อมูลประกอบการพิจารณา นอกจากนี้ยังมีงานวิจัยเกี่ยวกับการสร้างแบบจำลองการจัดหมวดหมู่สถานที่ท่องเที่ยวโดยใช้เทคนิคการเรียนรู้ของเครื่อง [5] ที่ใช้จัดหมวดหมู่สถานที่ท่องเที่ยวจากคำอธิบายสถานที่ท่องเที่ยว ซึ่งนำข้อมูลมาจากเว็บไซต์โลนลี่ แพลนเน็ตและได้นำเสนอแบบจำลองหลายประเภท แต่อย่างไรก็ตามงานวิจัยดังกล่าวรองรับกับการจัดหมวดหมู่สถานที่ท่องเที่ยวด้วยคำอธิบายสถานที่ท่องเที่ยวด้วยภาษาอังกฤษเท่านั้น

ดังนั้นพื้นฐานในการออกแบบและพัฒนาระบบจัดหมวดหมู่สถานที่ท่องเที่ยวด้วยคำอธิบายภาษาไทยแบบอัตโนมัติ นั้น มีการใช้ทฤษฎีพื้นฐานและเครื่องมือช่วยหลักๆ 3 ส่วน ได้แก่ 1) LexTo ซึ่งเป็นเครื่องมือที่ใช้สำหรับตัดแบ่งคำภาษาไทยจากคำอธิบายสถานที่ท่องเที่ยว เพื่อนำไปใช้ในการนับจำนวนคำสำคัญที่มีในคำอธิบาย 2) โปรแกรม Weka ซึ่งใช้ในการสร้างและเรียกใช้งานแบบจำลองที่ใช้ในการเรียนรู้ของเครื่องให้สามารถจัดหมวดหมู่ให้กับสถานที่ท่องเที่ยวตามคำอธิบายได้ และ 3) เทคนิคในการเรียนรู้ของเครื่องด้วยเทคนิคซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine: SVM)

2.1 LexTo (Thai Lexeme Tokenizer)

LexTo (<http://www.sansarn.com/lexto/>) เป็นเครื่องมือที่ใช้ในการแบ่งคำภาษาไทยที่ถูกพัฒนาขึ้นด้วยภาษาจาวา และมีการแจกจ่ายในรูปแบบของโอเพนซอร์ส ซึ่งมีต้นแบบมาจาก Lexitron เพื่อนำมาใช้ในการตัดแบ่งคำภาษาไทยโดยเฉพาะ โดย LexTo ใช้เทคนิคการแบ่งคำแบบคำที่รู้จักที่ยาวที่สุด (Longest Matching) หรือเรียกอีกอย่างหนึ่งว่าวิธีแบ่งคำแบบฐานพจนานุกรม (Dictionary-based) ซึ่งเป็นเทคนิคการแบ่งคำโดยเลือกคำที่ยาวที่สุดออกมาจากพจนานุกรม โดยพจนานุกรมที่กล่าวถึงนี้สามารถแก้ไข เพิ่มเติม และเปลี่ยนแปลงคำที่ต้องการตัดแบ่งได้อย่างอิสระ

หลักการทำงานของ Lexto คือ Lexto จะอ่านคำทั้งหมดจากพจนานุกรมเข้ามาในหน่วยความจำก่อน จากนั้นเมื่อมีการส่งข้อความที่ต้องการตัดแบ่งคำภาษาไทยเข้ามา Lexto จะนำคำในพจนานุกรมมาเทียบกับข้อความที่ถูกส่งเข้ามาโดยพิจารณาตัดแบ่งจากคำในพจนานุกรมที่ยาวที่สุด ดังตัวอย่างในภาพที่ 1 เมื่อต้องการแบ่งคำว่า “รถไฟฟ้าสีเหลือง” คำที่อยู่ในพจนานุกรมคือ “รถ” “รถไฟ” และ “รถไฟฟ้า” ซึ่งเป็นคำที่ยาวที่สุดจึงแบ่งคำแรกที่ “รถไฟฟ้า” จากนั้นจึงแบ่งคำถัดไป พบคำว่า “สี” และ “สีเหลือง” จึงแบ่งคำจากคำที่ยาวที่สุดซึ่งก็คือ “สีเหลือง” ดังนั้นเมื่อส่งความ “รถไฟฟ้าสีเหลือง” ให้กับ Lexto ก็จะได้คำ 2 คำออกมาคือ “รถไฟฟ้า” และ “สีเหลือง”



ภาพที่ 1: การตัดแบ่งคำภาษาไทยใน Lexto

2.2 โปรแกรม Weka

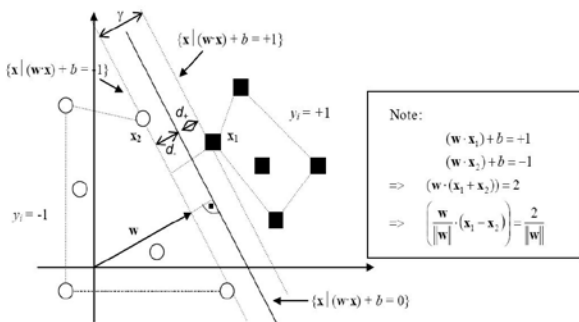
โปรแกรม Weka (Waikato Environment for Knowledge Analysis) เริ่มพัฒนามาตั้งแต่ปี 1997 โดยมหาวิทยาลัย Waikato ประเทศนิวซีแลนด์ ซึ่งเป็นฟรีแวร์ที่พัฒนาด้วยภาษาจาวา สามารถทำงานได้บนระบบปฏิบัติการหลักๆ ทุกตัว โดยเน้นกับ

งานทางด้าน การเรียนรู้ของเครื่อง (Machine Learning) และการทำเหมืองข้อมูล (Data Mining) อีกทั้งยังมีส่วนติดต่อสำหรับการเขียนโปรแกรม (API) ที่ผู้พัฒนาโปรแกรมต่างๆ สามารถติดต่อเรียกใช้งาน Weka ได้โดยตรง

Weka ให้บริการหน้าที่การทำงานพื้นฐานต่างๆ เกี่ยวกับการทำเหมืองข้อมูลอย่างมากมาย เช่น data processing, clustering, classification, regression, visualization และ feature selection รวมถึงยังเปิดช่องทางให้ผู้พัฒนาโปรแกรมสามารถนำอัลกอริทึมที่พัฒนาขึ้นมาเอง มาใช้ใน Weka ได้อีกด้วย

2.3 ซัพพอร์ตเวกเตอร์แมชชีน

หลักการของซัพพอร์ตเวกเตอร์แมชชีน คือ การสร้างไฮเปอร์เพลนที่เหมาะสมบนระนาบของข้อมูลตัวอย่าง (Training data) เพื่อแบ่งแยกกลุ่มข้อมูลที่แตกต่างกัน ในการสร้างไฮเปอร์เพลนที่เหมาะสม มีการนิยามระยะห่างระหว่างจุดของข้อมูลที่อยู่ใกล้กับไฮเปอร์เพลนมากที่สุดทั้งสองด้าน คือ d_+ และ d_- โดยระยะมารจิ้น (Margin) γ เกิดจากระยะ $d_+ + d_-$ ไฮเปอร์เพลนที่เหมาะสมคือไฮเปอร์เพลนที่มีค่ามารจิ้น γ กว้างที่สุดดังแสดงในภาพที่ 2 โดยข้อมูลตัวอย่างที่อยู่บนขอบของมารจิ้น γ จะถูกเรียกว่า *ซัพพอร์ตเวกเตอร์*



ภาพที่ 2: การแบ่งกลุ่มข้อมูล ■ และ ○ ด้วยไฮเปอร์เพลนใน SVM

จากภาพที่ 2 เป็นการแบ่งกลุ่มข้อมูลแบบ 2 กลุ่ม (Binary Classification Problem) กำหนดให้กลุ่มข้อมูลที่ใช้ในการฝึกสอน (Training Dataset) ประกอบด้วย l ตัวอย่าง (Samples) ซึ่งสามารถแสดงอยู่ในรูป $\{x_k, y_k\}, k = 1, \dots, l$ และ $x_k \in \mathcal{R}^n, y_k \in \{-1, +1\}$ โดย x_k จะเป็น อินพุตเวกเตอร์ ในขณะที่ y_k จะเป็นคลาสของข้อมูล (Class Label) หลักการของซัพพอร์ตเวกเตอร์แมชชีน คือ การสร้างไฮเปอร์เพลนที่

เหมาะสมบนระนาบของข้อมูลตัวอย่าง เพื่อแบ่งกลุ่มของข้อมูลทั้งสอง โดยไฮเปอร์เพลนดังกล่าวจะถูกกำหนดโดยพารามิเตอร์ (w, b) โดย w เป็นเวกเตอร์ที่ตั้งฉากกับไฮเปอร์เพลน และ b จะเป็นค่าคงที่ ซึ่งกำหนดตำแหน่งของเวกเตอร์ที่สัมพันธ์กับตำแหน่งดั้งเดิมในปริภูมิอินพุต (Input Space) สมการของไฮเปอร์เพลนแบบเชิงเส้น (Linear Hyperplane) จะถูกกำหนดโดยสมการ $(w \cdot x) + b = 0$ และเพื่อลดปัญหาในเรื่องของสเกล w และ b จะถูกกำหนดด้วยสมการ $|(w \cdot x) + b| = 1$ สำหรับจุดที่อยู่ใกล้ไฮเปอร์เพลนมากที่สุด ดังนั้นสมการของไฮเปอร์เพลนแสดงได้ดังสมการที่ (1)

$$y_i[(w \cdot x_i) + b] \geq 1 \quad \forall i \quad (1)$$

ดังที่กล่าวข้างต้นการฝึกสอนด้วยเทคนิค SVM จะเป็นการคำนวณหาไฮเปอร์เพลนที่มีค่ามารจิ้นกว้างที่สุดซึ่งสามารถหาได้จากการทำให้ค่า w มีค่าที่น้อยที่สุด โดยปัญหาดังกล่าวสามารถหาคำตอบได้โดยใช้วิธี Lagrange Multipliers ดังสมการที่ (2) โดยตัวแปร $\alpha_i \geq 0$ จะถูกเรียกว่า Positive Lagrange Multipliers

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^l \alpha_i y_i (x_i \cdot w + b) + \sum_{i=1}^l \alpha_i \quad (2)$$

จากสมการที่ (2) สิ่งที่ต้องการคือให้ค่าของ $L(w, b, \alpha)$ มีค่าน้อยที่สุด โดยเทียบกับค่าของ w และ b ในขณะเดียวกันค่าของ $L(w, b, \alpha)$ จะต้องมีค่ามากที่สุดเมื่อเทียบกับ $\alpha_i \geq 0$ (Dual Variables) โดยปัญหาดังกล่าวสามารถหาคำตอบได้โดยอาศัยวิธี Wolfe Dual ดังนั้นสมการที่ใช้เพื่อหาคำตอบของสมการที่ (2) สามารถแสดงได้ดังนี้

$$\text{Maximize } W(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \quad (3)$$

Subject to (1) $\sum_{i=1}^l \alpha_i y_i = 0$, and (2) $\alpha_i \geq 0$ for $i = 1, \dots, l$.

จากสมการที่ (3) ตัวอย่างข้อมูลที่ใช้ในการฝึกสอนที่ทำให้ค่าของ $\alpha_i > 0$ จะถูกเรียกว่า *ซัพพอร์ตเวกเตอร์* ซึ่งจะวางตัวอยู่บนขอบของมารจิ้น สำหรับข้อมูลตัวอย่างที่เหลือจะมีค่าของ $\alpha_i = 0$ สามารถจะถูกตัดทิ้งได้โดยไม่ก่อให้เกิดผลกระทบต่อไฮเปอร์เพลนที่สร้างขึ้น จากอัลกอริทึมที่กล่าวมาข้างต้น จะเหมาะสำหรับในกรณีที่กลุ่มข้อมูลตัวอย่างสามารถถูกแบ่งแยกได้ด้วยไฮเปอร์เพลนแบบเชิงเส้นเท่านั้น ดังนั้นเพื่อให้

อัลกอริธึมดังกล่าวสามารถแบ่งแยกกลุ่มข้อมูลที่มีลักษณะไม่เป็นเชิงเส้น (Nonlinear Dataset) จึงจำเป็นต้องแปลงกลุ่มข้อมูลตัวอย่างไปสู่ปริภูมิมิติที่สูงขึ้น (Higher Dimensional Space) เรียกว่าปริภูมิ Feature (Feature Space) โดยการแปลงดังกล่าวจะกระทำผ่านฟังก์ชันที่ไม่เป็นเชิงเส้นดังนี้ $\Phi: \mathcal{X}^n \rightarrow F$. กำหนดให้ Φ แสดงการแมพ (Mapping) จากปริภูมิของกลุ่มข้อมูลตัวอย่างไปสู่ปริภูมิ Feature โดยในที่นี้จะต้องรวมขั้นตอนดังกล่าวเข้ากับสมการที่ (3) ซึ่งสามารถทำได้โดยการแทนที่ทุกค่าของ \mathbf{x}_i ด้วย $\Phi(\mathbf{x}_i)$ โดยในที่นี้เราจะหลีกเลี่ยงการคำนวณค่า $\Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j)$ โดยตรงเนื่องจากเป็นขั้นตอนที่ต้องใช้เวลานานในคำนวณสูง แต่เราจะใช้เทคนิคของฟังก์ชันเคอร์เนล (Kernel Function) ในการคำนวณแทน โดยฟังก์ชันเคอร์เนลสามารถนิยามได้ดังนี้ $K(\mathbf{x}_i, \mathbf{x}_j) = \Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j)$. แทนที่เคอร์เนล K ลงในสมการที่ (3) ดังนั้นสามารถแสดงสมการซึ่งใช้ในการหาค่าของไฮเปอร์เพลนได้ดังสมการที่ (4) โดย C จะเป็นค่าคงที่เพื่อใช้ในการปรับหรือชดเชยระหว่างค่าผิดพลาดของการฝึกสอน

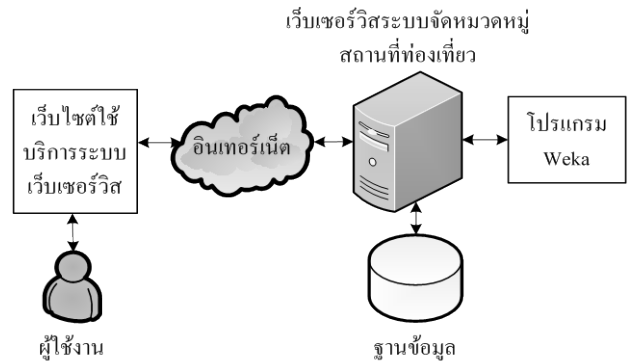
$$\text{Maximize } W(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_j) \quad (4)$$

$$\text{Subject to (1) } \sum_{i=1}^l \alpha_i y_i = 0, \text{ and (2) } 0 \leq \alpha_i \leq C \quad \forall i.$$

3. วิธีการดำเนินการวิจัย

ระบบจัดหมวดหมู่สถานที่ท่องเที่ยวในงานวิจัยนี้ ได้ถูกออกแบบให้บริการผ่านเว็บเซอร์วิส โดยมีหน้าที่หลัก 2 อย่างคือ 1) ให้บริการข้อมูลเกี่ยวกับ ชื่อ คำอธิบาย และหมวดหมู่สถานที่ท่องเที่ยวทั้งหมดที่มีอยู่ในฐานข้อมูลกับผู้ร้องขอ และ 2) ให้บริการในการรับข้อมูลของชื่อ และคำอธิบายสถานที่ท่องเที่ยวจากผู้ใช้งานทั่วไป เพื่อนำมาจัดหมวดหมู่ผ่านโปรแกรม Weka และเก็บข้อมูลทั้งหมดลงในฐานข้อมูลของระบบ ดังแสดงในภาพที่ 3

สำหรับการออกแบบและพัฒนาแบบจำลองการเรียนรู้ของเครื่อง สามารถแบ่งขั้นตอนสำคัญๆ ออกเป็น 2 ขั้นตอนคือการสกัดคุณลักษณะเด่นของหมวดหมู่สถานที่ท่องเที่ยว และการเลือกแบบจำลองการเรียนรู้ของเครื่องในการจัดหมวดหมู่สถานที่ท่องเที่ยว



ภาพที่ 3: โครงสร้างโดยรวมของระบบจัดหมวดหมู่สถานที่ท่องเที่ยว

3.1 การสกัดคุณลักษณะเด่นของหมวดหมู่สถานที่ท่องเที่ยว

ระบบจัดหมวดหมู่สถานที่ท่องเที่ยวต้นแบบนี้ ได้วางขอบเขตของสถานที่ท่องเที่ยวเฉพาะในจังหวัดภูเก็ต และมีความสามารถในการจำแนกหมวดหมู่สถานที่ท่องเที่ยวได้ทั้งหมด 4 หมวดหมู่ คือ วัด ทะเล ภูเขา และที่พัก

ดังนั้นขั้นตอนนี้ จึงเริ่มจากรวบรวมชื่อและคำอธิบายสถานที่ท่องเที่ยวในจังหวัดภูเก็ตจากอินเทอร์เน็ต หมวดหมู่ละ 10 สถานที่ แล้วนำมาผ่านโปรแกรมตัดคำภาษาไทย LexTo เพื่อนับจำนวนคำที่ถูกใช้ในคำอธิบายสถานที่ท่องเที่ยวมากที่สุด 15 อันดับแรกของแต่ละหมวดหมู่ ดังแสดงในตารางที่ 1

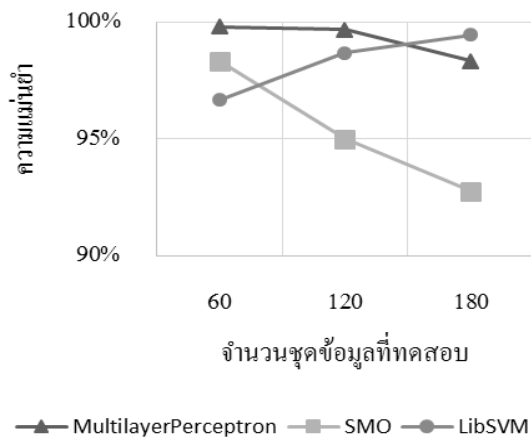
ตารางที่ 1: คำที่ถูกใช้บ่อยที่สุด 15 คำแรกในคำอธิบายสถานที่ท่องเที่ยว

อันดับ	วัด	ทะเล	ภูเขา	ที่พัก
1	วัด	หาด	ยอดเขา	รีสอร์ท
2	เทพ	ชายหาด	ป่า	สปา
3	พระ	กะตะ	ชม	ห้อง
4	เจ้าอาวาส	ทะเล	วิว	โรงแรม
5	หลวงพ่อ	หาดทราย	จุดชมวิว	บีช
6	บ้าน	เกาะ	ภูเขา	สิ่งอำนวยความสะดวก
7	ผูด	ภูเก็ต	รัง	ห้องพัก
8	ศรีลังกา	เต่า	ทัศนียภาพ	บริการ
9	ฉลอง	กิจกรรม	หอ	สระว่ายน้ำ
10	ชาวบ้าน	ป่าดง	กรมป่าไม้	ให้บริการ
11	ท้าว	แหลม	การอนุรักษ์	ไทย
12	นิมิต	คลื่น	ช้าง	เดอะ
13	บรม	น้ำ	ดินเขา	วิลล่า
14	พระทอง	ทราย	ทิวทัศน์	แกรนด์
15	พ่อ	ทางน้ำ	ป่าสงวน	ความสะดวก

3.2 การเลือกแบบจำลองการเรียนรู้ของเครื่อง

การเลือกแบบจำลองในการจัดหมวดหมู่ ถือเป็นหัวใจสำคัญของการเรียนรู้ของเครื่อง ผู้วิจัยจึงใช้ค่าสำคัญที่ถูกใช้บ่อยที่สุด 10 คำแรกของแต่ละหมวดหมู่ มาสร้างเป็นข้อมูลคุณลักษณะเด่นของสถานที่ท่องเที่ยวแต่ละสถานที่ ในการทดลองนี้ กลุ่มข้อมูลตัวอย่างของสถานที่ท่องเที่ยวที่ใช้ฝึกสอน และทดสอบความถูกต้องเป็นชุดข้อมูลเดียวกัน โดยได้ใช้ขนาดของชุดข้อมูล 3 ขนาด คือ 60, 120 และ 180 ข้อมูล และทดสอบกับแบบจำลองในการจัดหมวดหมู่ 3 แบบจำลอง คือ MultilayerPerceptron ที่เป็นแบบจำลองเครือข่ายประสาทเทียม และ ซัพพอร์ตเวกเตอร์แมชชีนที่ใช้วิธี SMO (Sequential minimal optimization) โดยทั้ง 2 แบบจำลองข้างต้นมีมาพร้อมกับ Weka และแบบจำลองที่ 3 คือ ซัพพอร์ตเวกเตอร์แมชชีนที่ติดตั้งเพิ่มเติมชื่อ LibSVM [6] ซึ่งใช้เคอร์เนลโดยปริยายคือฟังก์ชัน radial basis ได้ผลการทดลองดังภาพที่ 4

จากผลการทดลองจะเห็นว่า แบบจำลองที่ใช้เทคนิคของ MultilayerPerceptron ให้ความแม่นยำถึง 100% สำหรับการให้ข้อมูลฝึกสอนและทดสอบ ชุดเดียวกันทั้ง 60 ข้อมูล แต่ความแม่นยำเริ่มลดลงเมื่อมีจำนวนข้อมูลที่ฝึกสอนมากขึ้นเรื่อยๆ โดยมีความแม่นยำเหลือ 98.33% เพื่อมีชุดข้อมูลที่ฝึกสอนจำนวน 180 ข้อมูล สำหรับแบบจำลองที่ใช้เทคนิคซัพพอร์ตเวกเตอร์แมชชีน SMO ที่มีมาพร้อมกับ Weka นั้นให้ผลลัพธ์ที่ไม่น่าพึงพอใจ ซึ่งตรงกันข้ามกับแบบจำลองที่ใช้เทคนิคซัพพอร์ตเวกเตอร์แมชชีนของ LibSVM ที่ให้ความแม่นยำมากขึ้นเรื่อยๆ ตามจำนวนข้อมูลที่ใช้ในการฝึกสอน โดยให้ความแม่นยำถึง 99.44% สำหรับชุดข้อมูลที่ฝึกสอนจำนวน 180 ข้อมูล



ภาพที่ 4: ความแม่นยำของแบบจำลองแบบต่างๆ

4. ผลการดำเนินงาน

จากผลการทดสอบแบบจำลองประเภทต่างๆ จะเห็นได้ว่าแบบจำลองที่ใช้เทคนิคซัพพอร์ตเวกเตอร์แมชชีนของ LibSVM ให้ความแม่นยำมากที่สุด และเพื่อให้ระบบสามารถทำงานได้อย่างรวดเร็ว จึงแบ่งการทดลองเพิ่มเติมอีก 2 การทดลอง คือ การทดลองเพื่อวิเคราะห์จำนวนคำสำคัญที่ใช้เป็นคุณลักษณะเด่นในการจัดหมวดหมู่สถานที่ท่องเที่ยว ว่าถ้าลดจำนวนคำสำคัญออกไปจะมีผลกระทบต่อความแม่นยำและเวลาที่ใช้ในการจัดหมวดหมู่อย่างไร และการทดลองเพื่อหาความแม่นยำแบบละเอียดสำหรับหมวดหมู่แต่ละประเภท

4.1 ผลกระทบของแบบจำลองกับจำนวนคำสำคัญ

ในการทดลองนี้ได้นำคำที่ถูกพบบ่อยที่สุดของสถานที่ในแต่ละหมวดหมู่มาใช้เป็นลักษณะเด่นเพื่อสร้างแบบจำลอง โดยทดลองสร้างแบบจำลองทั้งหมด 3 แบบจำลอง จากคำที่พบบ่อยที่สุดของแต่ละหมวดหมู่จำนวน 5, 10 และ 15 คำแรก และในแต่ละแบบจำลองจะใช้ชุดข้อมูลฝึกจำนวน 5 ขนาด คือ 40, 60, 80, 100 และ 120 ข้อมูล เพื่อวัดเวลาที่ใช้ในการสร้างแบบจำลอง และความแม่นยำที่ใช้ในการจัดหมวดหมู่กับชุดข้อมูลทดสอบขนาด 40 ชุดข้อมูลที่ไม่ได้ถูกใช้ในขั้นตอนการฝึก ซึ่งได้ผลลัพธ์ดังแสดงในตารางที่ 2

จากการทดลองพบว่า การใช้คำสำคัญของแต่ละหมวดหมู่สถานที่ท่องเที่ยว 10 และ 15 คำ สำหรับจำนวนข้อมูลที่ฝึก 120 ข้อมูลนั้น ให้ความแม่นยำในการจัดหมวดหมู่ถึง 95% ซึ่งดีกว่าการใช้คำสำคัญเพียง 5 คำ ที่ให้ความแม่นยำเพียง 92.5% อย่างไรก็ตามจะเห็นว่าผลลัพธ์ของการใช้คำสำคัญ 10 คำนั้น ให้ผลลัพธ์เทียบเท่ากับการใช้คำสำคัญ 15 คำ และยังใช้เวลาในการสร้างแบบจำลองที่น้อยกว่า

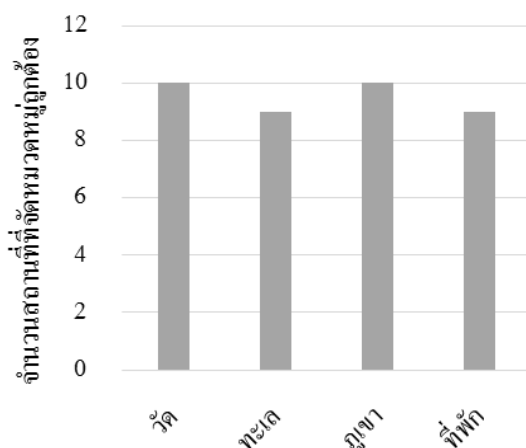
ตารางที่ 2: ความแม่นยำในการจัดหมวดหมู่ และเวลาในการสร้างแบบจำลองตามจำนวนคำสำคัญ

จำนวนข้อมูลที่ฝึก	5 คำสำคัญ		10 คำสำคัญ		15 คำสำคัญ	
	ความแม่นยำ	เวลา (ms)	ความแม่นยำ	เวลา (ms)	ความแม่นยำ	เวลา (ms)
40	77.5%	258	75.00%	304	72.5%	328
60	85.0%	314	87.5%	346	87.5%	386
80	87.5%	330	90.0%	388	90.0%	455
100	90.0%	357	92.5%	431	92.5%	500
120	92.5%	375	95.0%	488	95.0%	559

4.2 ความแม่นยำของแบบจำลองในแต่ละหมวดหมู่

จากผลการทดลองในหัวข้อที่ 4.1 พบว่าเมื่อใช้คำสำคัญจำนวน 10 คำสำคัญเพื่อเป็นคุณลักษณะเด่นในการสร้างแบบจำลอง และใช้จำนวนข้อมูลในการฝึก 120 ข้อมูลนั้น นอกจากจะทำให้ความแม่นยำที่สุดแล้ว ยังใช้เวลาในการสร้างแบบจำลองรวดเร็วกว่าการใช้คำสำคัญจำนวน 15 คำ ดังนั้นการทดลองนี้จะเป็นการทดสอบเพื่อดูว่าเมื่อใช้แบบจำลองที่เกิดจากการใช้คำสำคัญ 10 คำเป็นตัวกำหนดคุณลักษณะเด่นของหมวดหมู่สถานที่ท่องเที่ยวและใช้ข้อมูลจำนวน 120 ข้อมูลในการฝึกแบบจำลองแล้วนั้น เมื่อทดสอบกับข้อมูลที่ไม่ได้ใช้ในการฝึกหมวดหมู่ละ 10 ข้อมูลแล้ว ความแม่นยำในแต่ละหมวดหมู่แสดงดังภาพที่ 5

โดยการผลการทดลองพบว่าจากข้อมูลสถานที่ท่องเที่ยว 10 สถานที่ในแต่ละหมวดหมู่นั้น หมวดหมู่สถานที่ท่องเที่ยวประเภทวัดและภูเขามีความแม่นยำถึง 100% ซึ่งคาดว่าคำสำคัญที่ใช้กำหนดคุณลักษณะเด่นของสถานที่ท่องเที่ยวทั้ง 2 ประเภทนี้ค่อนข้างเด่นชัด เช่น วัด สำหรับหมวดหมู่ประเภทวัด และยอดเขา สำหรับหมวดหมู่ประเภทภูเขา แต่สำหรับหมวดหมู่ประเภทที่พักและทะเลนั้นมีการจัดหมวดหมู่ที่ผิดพลาดอยู่ประเภทละ 1 สถานที่ สำหรับหมวดหมู่ประเภททะเลคาดว่ามาจากคำสำคัญที่ใช้ในการกำหนดคุณลักษณะเด่น มีคำว่า กะตะ เป็นส่วนประกอบซึ่งเป็นคำเฉพาะเกินไป และสำหรับหมวดหมู่ประเภทที่พัก โดยทั่วไปจะมีคำอธิบายที่กล่าวถึงความใกล้ของที่พักกับภูเขาหรือชายหาดมากเกินไป ทำให้ถูกจัดหมวดหมู่ผิดประเภท



ภาพที่ 5: ความแม่นยำในการจัดหมวดหมู่สถานที่ท่องเที่ยว

5. สรุป

งานวิจัยนี้ได้ออกแบบและพัฒนาระบบจัดหมวดหมู่สถานที่ท่องเที่ยวด้วยคำอธิบายแบบอัตโนมัติผ่านเว็บเซอร์วิส โดยมีการประยุกต์นำความสามารถของเครื่องมือแบ่งคำภาษาไทย Lexto และโปรแกรม Weka มาใช้งานในการเรียนรู้ของเครื่องเพื่อช่วยเหลือการจัดหมวดหมู่สถานที่ท่องเที่ยวในจังหวัดภูเก็ต โดยมีหมวดหมู่สถานที่ท่องเที่ยวทั้งหมด 4 ประเภทได้แก่ วัด ทะเล ภูเขา และที่พัก จากการทดลองพบว่าการใช้แบบจำลองการเรียนรู้ของเครื่องจักรด้วยเทคนิคซ์พอร์ทเวกเตอร์แมชชีนด้วย LibSVM ให้ความแม่นยำในการจัดหมวดหมู่มากที่สุด อีกทั้งการใช้คำสำคัญของแต่ละหมวดหมู่สถานที่ท่องเที่ยวเพียง 10 คำ และจำนวนข้อมูลคำอธิบายของแต่ละสถานที่ที่ใช้ฝึกเพียง 120 ข้อมูล เพียงพอที่ทำให้แบบจำลองให้ความแม่นยำในการจัดหมวดหมู่ถึง 95%

เอกสารอ้างอิง

- [1] T.W. Miller and H.T. Nguyen., "Data and Text Mining: A Business Applications Approach", *New York: Pearson Prentice Hall Upper Saddle River*, 2005.
- [2] F. Sebastiani, "Text categorization", *Alessandro Zanasi (ed.), Text Mining and its Applications*, WIT Press, Southampton, UK, pp. 109-129, 2005.
- [3] A. Popescu, G. Grefenstette, and P-A. Moëllic, "Mining Tourist Information from User-Supplied Collections", *Proceedings of the 18th ACM conference on Information and knowledge management*. NY, USA, pp. 1713-1716, 2009.
- [4] A. Popescu, G. Grefenstette and H. Bouamor, "Mining a Multilingual Geographical Gazetteer from the Web", *Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology*, Washington DC., USA, vol 01, pp. 58-65, 2009
- [5] คมกิต ชัชวราภรณ์, ธรา อังสกุล และจิตติมนต์ อังสกุล, "แบบจำลองการจัดหมวดหมู่สถานที่ท่องเที่ยวโดยเทคนิคการเรียนรู้ของเครื่อง" *วารสารเทคโนโลยีสุรนารี ฉบับสังคมศาสตร์ ปีที่ 6 ฉบับที่ 2 (ธันวาคม 2555)* หน้า 35-58.
- [6] C.-C. Chang and C.-J. Lin., "LIBSVM : a library for support vector machines.", *ACM Transactions on Intelligent Systems and Technology*, vol 2, pp 27:1--27:27, 2011.

การพัฒนาเครื่องมือสำหรับให้ความเห็นประกอบและคัดกรองการเปลี่ยนแปลงลำดับเบส ของมนุษย์สำหรับกระบวนการ Next Generation Sequencing The Development of Next Generation Sequencing Tool for Annotation and Variant Filtrating of Human Nucleotide Alteration

กฤตบุญ ปัญจรัตนกร (Kittaboon Panjarattakorn)¹ อุบลวรรณ เขาวนกะกิจ (Ubolwan Chaovanakij)²

และ กิติ์สุชาติ พสุภา (Kitsuchart Pasupa)³

คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

¹ziscross@hotmail.com, ² evaaa.eve@gmail.com, ³ kitsuchart@it.kmitl.ac.th

บทคัดย่อ

บทความนี้นำเสนอการพัฒนาเครื่องมือสำหรับการให้ความเห็นประกอบและคัดกรองการเปลี่ยนแปลงลำดับเบสของมนุษย์สำหรับกระบวนการ Next Generation Sequencing ในการวิเคราะห์ข้อมูลเพื่อหาสาเหตุของการเกิดโรคทางพันธุกรรมในแต่ละครั้งจำเป็นต้องใช้เครื่องมือในการวิเคราะห์หลายตัวในการได้มาของผลลัพธ์ อีกทั้งใช้ระยะเวลาเนื่องจากข้อมูลที่ได้ออกมาจากเครื่องถอดรหัสพันธุกรรมมีปริมาณมหาศาล ปัจจุบันในขั้นตอนการกรองและการให้ข้อมูลประกอบยังต้องใช้ซอฟต์แวร์ที่มีประสิทธิภาพในการทำงาน ซึ่งมีค่าใช้จ่ายสูง อีกทั้ง ขั้นตอนและเทคนิคการวิเคราะห์ในแต่ละแห่งมีความแตกต่างกัน ขึ้นอยู่กับดุลยพินิจของแพทย์ ดังนั้นซอฟต์แวร์อาจจะถูกออกแบบมาไม่เหมาะสมกับแพทย์ที่ใช้ การพัฒนาเครื่องมือนี้จะช่วยให้การวิเคราะห์ข้อมูลทางพันธุกรรมมีความสะดวกรวดเร็ว และตรงกับกระบวนการวิเคราะห์ของผู้ใช้มากยิ่งขึ้น ซึ่งซอฟต์แวร์ที่นำเสนอนี้พัฒนาในรูปแบบเว็บแอปพลิเคชันโดยใช้ภาษา PHP

คำสำคัญ: เครื่องมือคัดกรอง เทคโนโลยีเอ็นจีเอส
เว็บแอปพลิเคชัน

Abstract

This paper presents the development of next generation sequencing tool for annotation and variant filtrating of human nucleotide alteration. Analysis of data to determine the cause of the genetic disease requires many tools to achieve the results and a lot of time due to huge amount of information obtained from the genetic decoder. Currently, the process of filtering and

annotation also needed an efficient software which is expensive. Each physician has different techniques and procedures, therefore the software might not be suitable for the physician. The development of this tool will enable the analysis of genetic information to be more convenient and faster. This software was developed as a web application and using PHP language.

Keywords: Filtrating Tool, NGS Technology, Web Application.

1. บทนำ

เทคโนโลยีที่ก้าวหน้าทางการแพทย์ในปัจจุบัน ทำให้เกิดการวินิจฉัยและรักษาโรคโดยอาศัยข้อมูลพื้นฐานทางพันธุกรรมมากขึ้น โดยการนำความรู้ทางด้านชีวสารสนเทศ (Bioinformatics) มาประยุกต์ใช้ คำว่ายีน (Genes) ดีเอ็นเอ (DNA) โครโมโซม (Chromosome) เริ่มเป็นที่รู้จักกันอย่างแพร่หลาย การค้นพบที่ประสบความสำเร็จอย่างมากคือการหาลำดับเบสจีโนมของมนุษย์โดยใช้เทคนิคที่เรียกว่า Next Generation Sequencing (NGS) ทำให้ได้ข้อมูลลำดับเบสจำนวนมหาศาล ตัวอย่างการใช้เทคนิคดังกล่าว เช่น การวิเคราะห์ลำดับเบสของยีนเพื่อวินิจฉัยโรคที่เกิดจากพันธุกรรมผิดปกติ การวิเคราะห์ยีนที่เป็นปัจจัยเสี่ยงในการแพ้ยา หรือการวิเคราะห์ปัจจัยเสี่ยงทางพันธุกรรมที่อาจก่อมะเร็ง [1] เป็นต้น ในปัจจุบันเครื่องมือสำหรับวิเคราะห์ข้อมูลยังมีข้อจำกัดอยู่มาก เช่น การสามารถเข้าถึงได้จากคอมพิวเตอร์เพียงเครื่องเดียว มีราคาแพง และใช้ฐานข้อมูลผู้ป่วยจากต่างประเทศเป็นส่วนใหญ่ ซึ่งมีลักษณะทางพันธุกรรมที่แตกต่างจากคนไทย ตัวอย่าง

เครื่องมือที่เป็นที่นิยมในปัจจุบันเช่น Ion Reporter [2] Golden Helix [3] และ Enlis [4] เป็นต้น

จากปัญหาดังกล่าว ทางผู้จัดทำจึงได้ร่วมมือกับภาควิชาอายุรศาสตร์และภาควิชาพยาธิวิทยา โรงพยาบาลรามธิบดี มหาวิทยาลัยมหิดล จัดทำเครื่องมือสำหรับการให้ความเห็นประกอบและคัดกรองการเปลี่ยนแปลงลำดับเบสของมนุษย์สำหรับกระบวนการ Next Generation sequencing ขึ้น เนื่องจากต้องการลดขั้นตอนการทำงาน ลดภาระการจดจำคำสั่งเพื่อช่วยให้แพทย์ใช้งานได้สะดวกมากขึ้น และเนื่องจากแต่ละศูนย์วิจัยมีความต้องการการใช้งานต่างกัน การพัฒนาเครื่องมือให้เหมาะสมจะสามารถช่วยระยะเวลาการทำงานได้ รวมถึงช่วยยกระดับหลักประกันสุขภาพของประเทศไทย ให้สามารถตรวจวินิจฉัยได้อย่างทั่วถึงในราคาที่ต่ำลง และยังสามารถคัดตัวอ่อนที่มีประสิทธิภาพในผู้ป่วยที่มีภาวะมีบุตรยากได้อีกด้วย

เนื้อหาในบทความส่วนที่ 2 จะกล่าวถึงวรรณกรรมที่เกี่ยวข้อง ส่วนที่ 3 กล่าวถึงวิธีการดำเนินงาน ซึ่งประกอบไปด้วยความต้องการของระบบ การออกแบบระบบ และภาพรวมของระบบ ผลการทดสอบการใช้งานจะปรากฏในส่วนที่ 4 และสรุปผลในส่วนที่ 5

2. วรรณกรรมที่เกี่ยวข้อง

2.1 โครโมโซม ยีน และดีเอ็นเอ

โครโมโซมเป็นที่อยู่ของดีเอ็นเอ (DNA) หรือสารพันธุกรรม รวมถึงหน่วยพันธุกรรมหรือยีน โครโมโซมทำหน้าที่ควบคุมและถ่ายทอดลักษณะทางพันธุกรรมต่างๆ ของสิ่งมีชีวิตเช่น ลักษณะของดวงตา เส้นผม สีผิวและเพศ หากเปรียบเทียบให้เข้าใจง่าย ดีเอ็นเอเปรียบเสมือนตัวหนังสือ ก.ไก่ ถึง ฮ.นกฮูก เป็นตัวๆ สลับสับเปลี่ยนกันเป็นประโยคให้อ่านรู้เรื่อง และรวมกันเป็นบทต่างๆ ในหนังสือซึ่งอาจมีความยาวไม่เท่ากัน บทต่างๆ นั้นเปรียบเสมือนยีน (Gene) ที่ประกอบไปด้วยลำดับต่างๆ ของดีเอ็นเอที่สามารถตีความ ถอดรหัสได้ แต่ละบทก็สามารถรวมเป็นหนังสือหนึ่งเล่มที่เปรียบได้กับโครโมโซม (Chromosome) ในร่างกายมนุษย์ เป็นต้น

2.2 โรคทางพันธุกรรม

โรคพันธุกรรม (Genetic Disorder) เป็นโรคที่เกิดจากความผิดปกติของดีเอ็นเอในแต่ละคน โดยความผิดปกตินั้นอาจเกิดได้ตั้งแต่ความผิดปกติของยีนเดี่ยว ไปจนถึงการเพิ่มหรือการ

ขาดหายไปของโครโมโซม ซึ่งโรคทางพันธุกรรมอาจมีสาเหตุมาจากการถ่ายทอดทางพันธุกรรมของพ่อและแม่ หากยีนหรือโครโมโซมของพ่อและแม่มีความผิดปกติแฝงอยู่ โรคทางพันธุกรรมเป็นโรคที่พบได้น้อย อาจพบผู้ป่วยเพียงหนึ่งในหลายพันคนหรือหลายล้านคน แต่เมื่อนำมารวมกันแล้ว การเกิดโรคเกือบจะเป็นครึ่งหนึ่งของการเกิดโรคในคนทั้งหมด

2.3 สนิปส์

สนิปส์ หรือ SNPs (Single Nucleotide Polymorphisms) เป็นความแตกต่างหรือความหลากหลายทางพันธุกรรมที่พบมากที่สุดในหมู่ประชากรมนุษย์ เกิดจากการเปลี่ยนแปลงลำดับเบสบนสายดีเอ็นเอเพียงตำแหน่งเดียวและก่อให้เกิดผลทางกายภาพ

ในร่างกายของคนเรามีโครโมโซมทั้งหมด 23 โครโมโซม ซึ่งประกอบกันเป็นหนึ่งชุดของมนุษย์เรียกว่าจีโนม ในจีโนมมนุษย์นั้นประกอบด้วยคู่เบสถึงสามพันล้านคู่ ถ้าหากเรานำลำดับเบสของแต่ละคนมาเปรียบเทียบกันจะพบว่ามีค่าเหมือนกันถึง 99.9% ส่วนที่เหลืออีก 0.1% นั้นเป็นบริเวณที่ทำให้เราแตกต่างจากคนอื่น ซึ่งบริเวณตรงนี้เองที่ถูกอธิบายด้วยสนิปส์ (ภาพที่ 1) อย่างไรก็ตามการเปลี่ยนแปลงลำดับเบสในแต่ละจุดไม่ได้ถูกเรียกว่าสนิปส์เสมอไป ความแตกต่างที่เกิดขึ้นจัดเป็น สนิปส์ก็ต่อเมื่อพบในประชากรมากกว่า 1% แต่หากน้อยกว่า 1% จะถือว่าเป็นการกลายพันธุ์เฉพาะจุด (Point Mutation) เท่านั้น



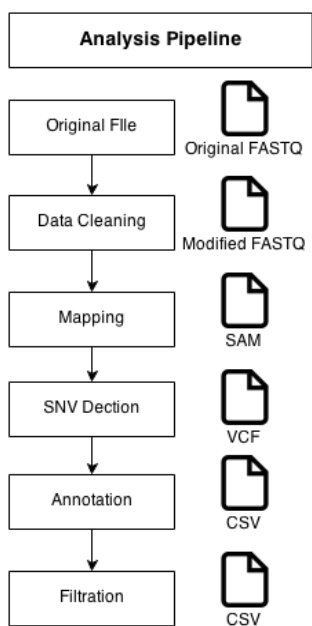
ภาพที่ 1: ตัวอย่างการเกิดสนิปส์

2.4 ขั้นตอนของกระบวนการเอ็นจีเอส

ขั้นแรกเริ่มจากการหาลำดับเบสด้วยเครื่องอ่านลำดับเบส ซึ่งจะได้อไฟล์ .fastq ที่ประกอบด้วยข้อมูลดีเอ็นเอของคนไข้ จากนั้นจึงกรองข้อมูลที่มีคุณภาพต่ำทิ้ง (Data Cleaning) ตามด้วยการเปรียบเทียบข้อมูลโดยการนำลำดับเบสของผู้ป่วยมา

เปรียบเทียบกับประชากรปกติ (Mapping) โดยในขั้นตอนนี้จะได้ไฟล์ .sam ซึ่งผ่านการจัดเรียงเรียบร้อยแล้ว โดยจะบอกรายละเอียดต่าง ๆ ที่สำคัญเช่น Reference Sequence, Read Group เป็นต้น และจึงตรวจดูว่าโปรตีนในตำแหน่งที่ต่างกันั้นเป็นสลับหรือการกลายพันธุ์ (SNV Detection) ในขั้นตอนนี้จะได้ผลลัพธ์เป็นไฟล์ .vcf ซึ่งจะระบุรายละเอียดมากขึ้น เช่น โครโมโซม ตำแหน่งบนโครโมโซม โปรตีน เป็นต้น และเลือกเฉพาะที่เป็นสลับมาพิจารณาให้ความเห็นประกอบ โดยอาจนำไปเปรียบเทียบกับฐานข้อมูลต่างๆ (Annotation) เช่น 1000Genomes [5] dbSNP [6] เพื่อให้ได้รายละเอียดที่มากขึ้นซึ่งจะได้ไฟล์ .csv หลังจากนั้นจึงคัดกรองยีนที่ต้องสงสัยด้วยการใช้การคัดกรองแบบต่างๆ (Filtration) ต่อไป

ขั้นตอนการทำงานของกระบวนการเอ็นจีเอส หรือ Next Generation Sequencing (NGS) สามารถอธิบายได้ดังภาพที่ 2



ภาพที่ 2: ขั้นตอนการทำงานของกระบวนการเอ็นจีเอส

3. วิธีการดำเนินงาน

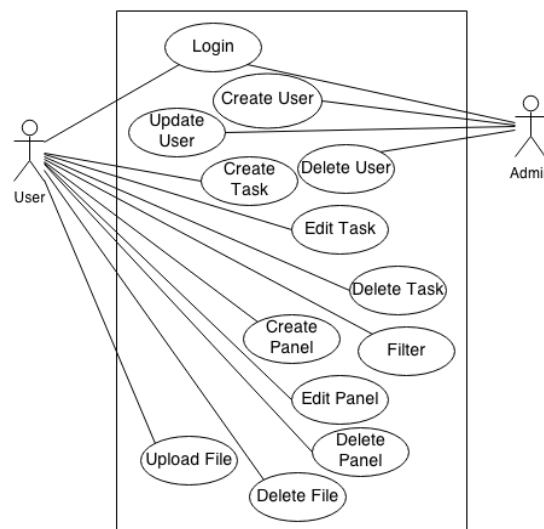
3.1 ปัญหาและความต้องการระบบ

ปัจจุบันทางภาควิชาได้ใช้โปรแกรมสำเร็จรูป Golden Helix ช่วยในการวิเคราะห์ข้อมูลซึ่งทำงานในส่วน Filtration ซึ่งการใช้งานจำเป็นต้องดาวน์โหลดข้อมูลจำนวนมากมาเก็บไว้ที่เครื่องคอมพิวเตอร์ และทำการคัดกรองต่าง ๆ ตามกระบวนการของแพทย์ ซึ่งเครื่องมือที่ใช้อยู่ไม่สามารถค้นหาข้อมูลอื่น ๆ

เพิ่มเติมนอกเหนือจากที่มีอยู่ได้โดยตรง เช่น ในการค้นหารายละเอียดเพิ่มเติมของยีนที่เกี่ยวข้อง ไม่สามารถคลิกบนชื่อยีนและเชื่อมต่อไปยังฐานข้อมูลภายนอกได้ ต้องนำชื่อยีนไปค้นหาเองอีกต่อหนึ่ง

จากการสอบถามความต้องการจากผู้ใช้งานพบว่าความต้องการที่จะพัฒนาระบบแทนที่ระบบปัจจุบันที่ใช้อยู่ โดยมีความต้องการเพิ่มเติมคือ ตัวระบบ สามารถคลิกที่ยีนที่ต้องการและเชื่อมต่อไปยังแหล่งข้อมูลภายนอกได้ทันที เช่น ฐานข้อมูล OMIM [7] ซึ่งเป็นฐานข้อมูลของยีนและอาการของโรคที่เกิดจากพันธุกรรม

ระบบที่พัฒนาขึ้นแบ่งผู้ใช้งานออกเป็นสองระดับคือ (1) ผู้ใช้งานทั่วไป สามารถเข้าสู่ระบบ สร้าง/แก้ไข/ลบการทำงานที่ต้องการ สามารถทำการคัดกรองข้อมูลเบื้องต้นและปรับแต่งตัวคัดกรองได้ สามารถสร้าง/แก้ไข/ลบรายชื่อยีนกลุ่มโรค รวมถึงสามารถอัปโหลดไฟล์เข้าระบบและลบไฟล์ที่ไม่ต้องการได้ (2) ผู้ดูแลระบบ สามารถเข้าสู่ระบบเพื่อเพิ่ม/แก้ไข/ลบผู้ใช้งานได้ โดยยูสเคสไดอะแกรมจะแสดงดังภาพที่ 3

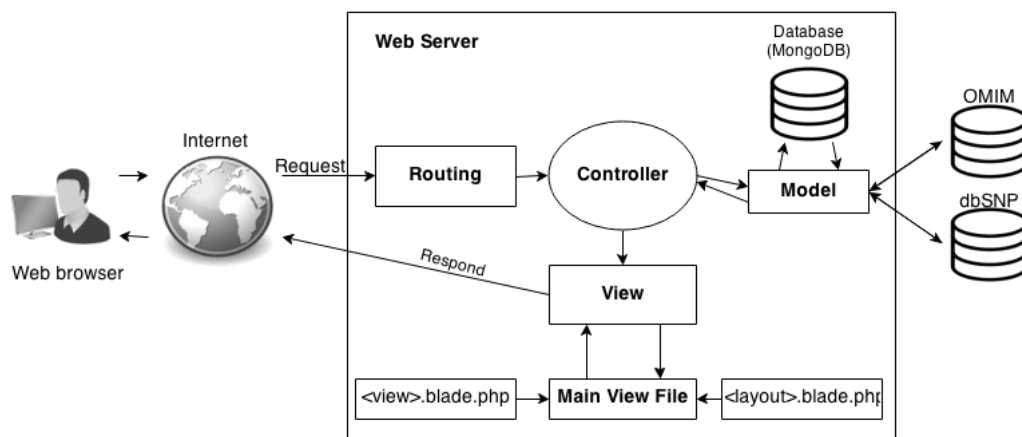


ภาพที่ 3: ยูสเคสไดอะแกรม

3.2 การออกแบบระบบ

ระบบที่นำเสนอแบ่งเป็น 2 ส่วนคือ Web Application และ Web Server โดยมีรายละเอียดดังนี้

Web Application พัฒนาโดยใช้ภาษา PHP เนื่องจากทำงานฝั่งเซิร์ฟเวอร์ (Server-side Script Language) สามารถฝังคำสั่ง PHP ไว้ในเว็บเพจร่วมกับ HTML Tag ได้ ช่วยให้เว็บไซต์มีความปลอดภัยมากขึ้นและประมวลผลได้รวดเร็ว



ภาพที่ 4: การออกแบบระบบโดยใช้ Laravel Framework

Web Server ออกแบบโดยใช้หลักสถาปัตยกรรมซอฟต์แวร์แบบ MVC (Model-View-Control) เพื่อให้ง่ายต่อการแยกส่วนพัฒนาและสามารถนำกลับมาใช้ใหม่ได้ โดยเลือกใช้ Laravel Framework [8] ช่วยในการพัฒนาเนื่องจากโค้ดมีความเรียบร้อย ตรวจสอบแก้ไขได้ง่าย มีความยืดหยุ่นสูงและเป็นเอกเทศต่อกัน รายละเอียดของ Laravel Framework จะคล้ายกับ PHP Framework ทั่วไป ใช้ Model ในการติดต่อกับฐานข้อมูล ใช้ Controller ในการรับข้อมูล (Input) จากผู้ใช้และประสานงานระหว่าง Model และ View และแสดงผลทั้งหมดด้วย View แต่ต่างกับที่ Laravel มีจุดเด่นพิเศษคือ Blade Template ซึ่งภายในไฟล์จะแบ่งออกเป็นส่วน ๆ ช่วยให้สะดวกต่อการแบ่งส่วนการแสดงผล และสามารถสร้างไฟล์ต้นแบบ (Template) หน้าเว็บขึ้นเองได้

จากภาพที่ 4 ขั้นตอนการทำงานเริ่มจากผู้ใช้ทำการเชื่อมต่ออินเทอร์เน็ตและเข้าสู่ระบบการทำงาน โดยระบบจะส่ง Request การทำงานไปยังส่วน Routing ซึ่งจะประมวลผล URL ที่ผู้ใช้ต้องการและเรียกใช้ Controller ขึ้นมา ซึ่ง Controller จะประสานงานร่วมกับ Model ในการดึงข้อมูลจากฐานข้อมูลมาแสดงผล แล้วจึงส่งค่าที่ได้ไปให้ View ในการแสดงผลลัพธ์ โดยระบบสามารถเชื่อมต่อกับฐานข้อมูลจากภายนอกได้ เช่น OMIM, dbSNP โดยการส่ง GET ค่าผ่าน URL

ฐานข้อมูลเลือกใช้ MongoDB [9] ซึ่งไม่มีความจำเป็นต้องอาศัยความสัมพันธ์ระหว่างเอ็นติตีภายในฐานข้อมูล อีกทั้งยังสามารถเก็บข้อมูลในลักษณะ Semi-structure ในรูปแบบ JSON ซึ่งเหมาะสมกับลักษณะข้อมูลในงานนี้ที่ข้อมูลมีโครงสร้างไม่แน่นอนอีกด้วย

3.3 การทำงานของระบบ

เมื่อเริ่มต้นการใช้งาน ผู้ใช้งานต้องลงชื่อเข้าใช้ระบบเป็นอันดับแรก และจะเข้าสู่หน้าจอหลักสำหรับการใช้งาน จากนั้นจะต้องทำการอัปโหลดไฟล์ Sample เพื่อเริ่มต้นการทำงาน และทำการสร้างงานใหม่และ Import ไฟล์ที่ต้องการ จากนั้นจึงทำการคัดกรองจนได้ผลลัพธ์ในที่สุด โดยระบบมีหน้าจอกการทำงานต่างๆ ดังนี้

1. การเข้าสู่ระบบ

เป็นส่วนแรกที่ผู้ใช้งานต้องลงชื่อเข้าใช้ระบบเพื่อที่ระบบจะประมวลผลและแสดงผลหน้าจอให้กับผู้ใช้ได้ตามสถานะที่ผู้ใช้เป็นอยู่ ดังภาพที่ 5

2. หน้าจอหลักการใช้งาน

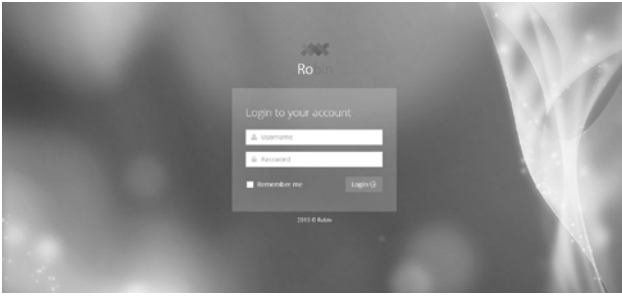
จะแสดงจำนวนไฟล์ทั้งหมดที่มีในระบบและจำนวนงานทั้งหมด พร้อมแสดงรายชื่องานที่มีการอัปเดตล่าสุด ดังภาพที่ 6

3. การอัปโหลดไฟล์

ผู้ใช้สามารถอัปโหลดไฟล์ที่ต้องการได้จากเมนูนี้ และสามารถลบไฟล์ที่ไม่ต้องการออกได้ ดังภาพที่ 7

4. การสร้างการทำงาน

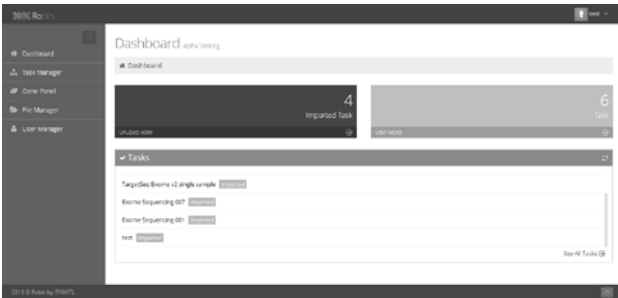
สำหรับผู้ใช้ที่ต้องการสร้างการทำงานใหม่ โดยสามารถแบ่งเป็นการวิเคราะห์แบบ Single, Paired หรือ Trio (เปรียบเทียบแบบพ่อ-แม่-ลูก) โดยนำเข้าไฟล์จากที่ได้ทำการอัปโหลดไว้ สามารถแก้ไข และลบการทำงานได้ ดังภาพที่ 8



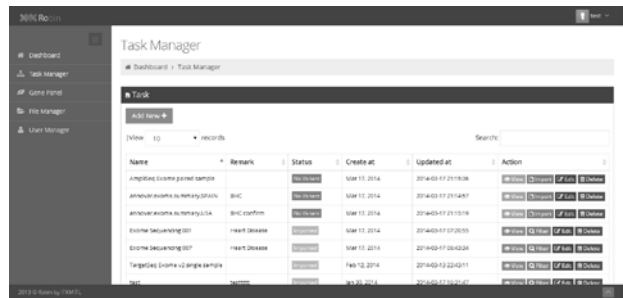
ภาพที่ 5: หน้าจอของระบบส่วนการเข้าสู่ระบบ



ภาพที่ 7: การแสดงผลการอัปโหลดไฟล์



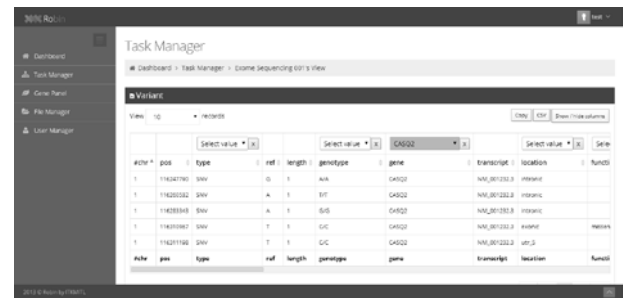
ภาพที่ 6: หน้าจอหลัก



ภาพที่ 8: หน้าจอของระบบส่วนการสร้างการทำงาน

5. การคัดกรองเบื้องต้น

จะแสดงรายละเอียดต่าง ๆ ของ Sample เช่น ตำแหน่งบนโครโมโซม จีโนไทป์ ยีน เป็นต้น ซึ่งเมื่อคลิกบนชื่อยีนจะสามารถเชื่อมต่อกับฐานข้อมูล OMIM เพื่อดูรายละเอียดที่เกี่ยวข้องได้ และสามารถคัดกรองข้อมูลต่าง ๆ เช่น Exonic Function, Genotype, SIFT, PolyPhen ดังภาพที่ 9 โดยขั้นตอนการคัดกรองข้อมูลต่าง ๆ นั้นขึ้นอยู่กับคุณสมบัติของแพทย์



ภาพที่ 9: หน้าจอของระบบส่วนการคัดกรองเบื้องต้น

4. ผลการดำเนินงาน

4.1 การทดสอบระบบ

ทดสอบเบื้องต้นกับครอบครัวที่มีผู้ป่วยเป็นโรคหัวใจที่ถ่ายทอดทางพันธุกรรมพบว่าสามารถค้นหายีนที่ก่อให้เกิดโรคได้ โดยมีขั้นตอนการคัดกรองเบื้องต้นเริ่มจาก ตรวจสอบลักษณะการถ่ายทอดของโรคทางพันธุกรรม จากนั้นจึงคัดกรองข้อมูล เช่น กรอง Exonic Function กรอง Genotype และข้อมูลอื่น ๆ ตามคุณสมบัติของแพทย์ เพื่อให้ได้ Candidate Gene ที่เหลือน้อยที่สุด ซึ่งระบบสามารถให้ความเห็นประกอบผ่านฐานข้อมูล OMIM, dbSNP ได้ หลังจากนั้นจึงนำ Candidate Gene ที่ได้ตรวจวิเคราะห์ด้วยการทำ Sanger Sequencing [10] เพื่อทำการยืนยันผลต่อไป

4.2 ผลการเปรียบเทียบการทำงานกับระบบเดิม

ผลการเปรียบเทียบความแตกต่างสามารถแสดงได้ดังตารางที่ 1 ระบบที่นำเสนอมีลักษณะเป็นเว็บแอปพลิเคชันทำให้สามารถใช้งานง่ายและสะดวกมากกว่าเนื่องจากใช้งานผ่านอินเทอร์เน็ต เพียงแค่มีเว็บเบราว์เซอร์ซึ่งเป็นสิ่งพื้นฐานในคอมพิวเตอร์ปัจจุบันเกือบจะทุกเครื่องและผู้ใช้มีความคุ้นเคยอยู่แล้วทำให้ลดค่าใช้จ่ายในการฝึกอบรมและยังสามารถใช้งานได้หลากหลายแพลตฟอร์มทั้ง Windows, Linux และ Mac ในขณะที่ระบบเดิมสามารถทำงานได้เฉพาะเครื่องที่มีการติดตั้งโปรแกรมเท่านั้น หากคำนึงถึงเรื่องการเก็บข้อมูลและการประมวลผลข้อมูล ตัวระบบที่นำเสนอก็มีข้อดีกว่ามากเนื่องจากมีการประมวลผลและจัดเก็บข้อมูลแบบออนไลน์ ข้อมูลถูกจัดเก็บที่เดียวทำให้ง่ายต่อการจัดการและไม่เกิดความซ้ำซ้อน

แต่อาจมีข้อเสียในเรื่องความเร็วในการประมวลผลเพราะต้องประมวลผลผ่านอินเทอร์เน็ตอาจทำให้ไม่รวดเร็วเท่าระบบเดิมที่สามารถใช้ทรัพยากรบนเครื่องคอมพิวเตอร์ได้อย่างเต็มที่เมื่อพิจารณาถึงการวิเคราะห์ข้อมูลและการกรองข้อมูล ทั้งสองระบบสามารถวิเคราะห์ข้อมูล Sample ได้มากกว่า 1 ราย และกรองข้อมูลโดยใช้ซินกลุ่มโรคได้ แต่ระบบที่นำเสนอมีข้อดีกว่าที่สามารถให้ข้อมูลเพิ่มเติมผ่านฐานข้อมูล OMIM โดยคลิกที่ชื่อยีนที่ต้องการ ซึ่งตรงกับความต้องการของผู้ใช้งาน ทั้งระบบใหม่ ทั้งยังสามารถกรองข้อมูลได้แบบ Interactive กับผู้ใช้สามารถใช้งานได้ง่ายและสะดวกรวดเร็ว รวมทั้งช่วยลดค่าใช้จ่ายได้อีกด้วย

ตารางที่ 1: เปรียบเทียบคุณสมบัติ

คุณสมบัติ	Golden Helix	ระบบที่พัฒนา
ลักษณะ โปรแกรม	วินโดวส์แอปพลิเคชัน	เว็บแอปพลิเคชัน
การประมวลผลและการเก็บข้อมูล	ออฟไลน์	ออนไลน์
วิเคราะห์ Sample ได้คราวละมากกว่า 1 ราย	✓	✓
กรองข้อมูลโดยอาศัยซินกลุ่มโรค	✓	✓
การเพิ่มตัวกรองข้อมูลเอง	✓	✓
การค้นหาข้อมูลจากภายนอก	✗	✓
ค่าลิขสิทธิ์	100,000 บาท/ปี	ไม่เสียค่าใช้จ่าย

5. สรุปผล

โดยสรุปการพัฒนาเครื่องมือสำหรับการให้ความเห็นประกอบและคัดกรองการเปลี่ยนแปลงลำดับเบสของมนุษย์สำหรับกระบวนการ Next Generation Sequencing โดยใช้ภาษา PHP สามารถช่วยให้การวิเคราะห์ข้อมูลมีความสะดวกมากขึ้นลดความซ้ำซ้อนในการจัดเก็บข้อมูล เชื่อมต่อกับฐานข้อมูลภายนอกต่างๆ เช่น 1000Genomes, dbSNP, OMIM เพื่อช่วยเหลือการวิเคราะห์ของแพทย์ อีกทั้งยังสามารถปรับแต่งตัวคัดกรองให้ตรงกับความต้องการได้ และช่วยลดค่าใช้จ่าย

กิตติกรรมประกาศ

โครงการนี้ได้รับการสนับสนุนจาก โครงการเกศขพันธุศาสตร์ ภายใต้ความร่วมมือระหว่างศูนย์ความเป็นเลิศด้านชีววิทยาศาสตร์ (องค์การมหาชน) และมหาวิทยาลัยมหิดล และศูนย์จีโนมทางการแพทย์ คณะแพทยศาสตร์ โรงพยาบาลรามาธิบดี

เอกสารอ้างอิง

- [1] A. Wilantho, O. Praditsup, W. Charoenchim, S. Kulawonganchai, A. Assawamakin, and S. Tongsimma “Next Generation Sequencing (NGS) Technologies and their Applications in Omics-research” Thai Journal of Genetics, vol, 5, no. 2, pp. 104-129.
- [2] Ion Reporter [Software]. Carlsbad, CA: Life Technologies Corporation, Available at <http://ionreporter.lifetechnologies.com/ir>. [Access: 30 September 2013]
- [3] SNP & Variation Suite (Version 7.x) [Software]. Bozeman, MT: Golden Helix, Inc., Available at <http://www.goldenhelix.com>. [Access: 30 September 2013]
- [4] Enlis [Software]. Berkeley, CA: Enlis, Llc., Available at <http://www.enlis.com>. [Access: 30 September 2013]
- [5] 1000 Genomes Project Consortium, “A map of human genome variation from population-scale sequencing” *Nature*, vol. 467, pp. 1061-1073, 2010.
- [6] S. Sherry, M. Ward, M. Kholodov, J. Baker, L. Phan, E. M. Smigielski, and K. Sirotkin, “dbSNP: the NCBI database of genetic variation” *Nucleic Acids Research*, vol. 29, no.1, pp. 308-311, 2001.
- [7] J. Amberger, C. Bocchini, A. F. Scott and A. Hamosh. “McKusick’s Online Mendelian Inheritance in Man (OMIM®)”. *Nucleic Acids Research*, vol. 37, pp.D793-D796, 2008.
- [8] Taylor Otwell. Laravel – The PHP framework for web artisans [Internet], Available at: <http://laravel.com/>. [Access: 10 October 2013]
- [9] MongoDB, Inc. MongoDB, Available at: <http://www.mongodb.org/>. [Access: 10 October 2013]
- [10] F. Sanger, S. Nicklen, and A. R. Coulson, “DNA Sequencing with chain-terminating inhibitors”, *Proceedings of the National Academy of Sciences of the United States of America*, vol. 74, no. 12, pp. 5463-5467, 1977.

การสำรวจและสร้างแผนที่ความพร้อมใช้งานของเครือข่ายด้วยคราวด์ซอร์ซซิง

Surveying and Mapping Network Availability using Crowdsourcing

ธภัทร หมู่ภัทร โรจน์ (Thaphat Mupattararot)¹ และ วีระ เหมืองสิน (Veera Muangsinsin)²

ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

¹thaphat.mu@hotmail.com, ²vmuangsinsinveera.m@chula.ac.th

บทคัดย่อ

แผนที่แสดงพื้นที่ให้บริการเครือข่ายไร้สายโดยทั่วไปได้มาจากการคำนวณโดยใช้ตำแหน่งอุปกรณ์รับส่งสัญญาณและรัศมีที่สามารถรับส่งสัญญาณได้ แผนที่ที่ได้จึงมีความแม่นยำน้อย แผนที่ที่ได้จากการสำรวจประสิทธิภาพของเครือข่ายจากพื้นที่จริงมีความแม่นยำสูงกว่าแต่ต้องใช้กำลังคนและอุปกรณ์จำนวนมากและมีค่าใช้จ่ายสูง งานวิจัยนี้จึงนำเสนอระบบสร้างแผนที่แสดงระดับสัญญาณทั้งของเครือข่ายโทรศัพท์เคลื่อนที่และ Wi-Fi จากการเก็บข้อมูลสัญญาณโดยใช้สมาร์ตโฟนจำนวนมากในลักษณะของการใช้พลังมวลชน (Crowdsourcing) ซึ่งสามารถเก็บข้อมูลได้ทั้งภายในและภายนอกอาคาร และแม้ในบริเวณที่ไม่สามารถทำการเชื่อมต่ออินเทอร์เน็ต และใช้วิธีการคำนวณความน่าจะเป็นในการคำนวณหาความแม่นยำของผลลัพธ์ที่ได้เพื่อการแสดงผลให้ดียิ่งขึ้น

คำสำคัญ: แผนที่สัญญาณ พลังมวลชน ความพร้อมใช้ของสัญญาณไร้สาย

Abstract

A map showing service area for wireless network is normally created from calculation based on the positions of signal transmission devices and their transmission range. Such a map has low accuracy. A map created from field survey of effective network connection has higher accuracy but requires man power, devices and expenses. In this paper, we propose a system to build maps showing signal levels for mobile and Wi-Fi networks by using smartphones to collect signal data in a crowdsourcing fashion. The data can be collected indoor and outdoor and even when the device cannot connect to the internet.

Moreover, we use a probabilistic method to calculate the accuracy of the obtained results.

Keyword: Network Availability, Wireless Network, Signal Map, Crowdsourcing.

1. ความเป็นมาและความสำคัญของปัญหา

ในปัจจุบัน การสื่อสารแบบไร้สายหรือ Wireless Communication ผ่านสมาร์ตโฟนนั้น เป็นไปอย่างแพร่หลาย เพราะสามารถใช้บริการจากเครือข่ายโทรศัพท์เคลื่อนที่ (Mobile Network) และเครือข่ายไวไฟ (Wi-Fi Network) จากผู้ให้บริการต่างๆ แต่พื้นที่การให้บริการอาจจะไม่ครอบคลุมได้ตามต้องการ ดังนั้น ถ้าผู้ใช้งานสามารถทราบถึงความแรงของสัญญาณหรือความพร้อมใช้งานของเครือข่าย (Network Availability) ของเครือข่ายต่างๆ ในแต่ละพื้นที่ก็จะมีประโยชน์อย่างยิ่ง โดยทั่วไปแผนที่แสดงพื้นที่ให้บริการเครือข่ายไร้สายได้มาจากการคำนวณโดยใช้ตำแหน่งอุปกรณ์รับส่งสัญญาณและรัศมีที่สามารถรับส่งสัญญาณได้ แผนที่ที่ได้จึงมีความแม่นยำน้อย แผนที่ที่ได้จากการสำรวจประสิทธิภาพของเครือข่ายจากพื้นที่จริงมีความแม่นยำสูงกว่า อย่างไรก็ตาม การสำรวจคุณภาพสัญญาณโดยทั่วไปต้องใช้อุปกรณ์สำหรับวัดสัญญาณโดยเฉพาะ และใช้บุคลากรเดินหรือขับรถสำรวจ ซึ่งใช้ระยะเวลานาน โดยเฉพาะกับพื้นที่ขนาดใหญ่และเสียค่าใช้จ่ายสูงในการดำเนินการ

แต่ในปัจจุบัน สมาร์ตโฟนก็มีความสามารถที่จะทำงานเป็นอุปกรณ์วัดสัญญาณเครือข่าย และยังสามารถระบุพิกัดตำแหน่งได้ด้วย ดังนั้นผู้วิจัยจึงมีความเห็นว่า หากเราสามารถเก็บข้อมูลสัญญาณจากสมาร์ตโฟนที่มีการใช้งานกันอย่างแพร่หลาย โดยใช้พลังมวลชน หรือ คราวด์ซอร์ซซิง (Crowdsourcing) ก็จะสามารถสำรวจข้อมูลเกี่ยวกับเครือข่ายโดยผู้ที่ใช้งานจริง และ

เป็นการลดค่าใช้จ่ายและเวลา ข้อมูลที่ได้สามารถนำมาสร้างเป็นแผนที่ความแรงของสัญญาณและความพร้อมใช้งานของเครือข่ายต่างๆ ทั้งนี้ผู้ใช้ควรสามารถระบุตำแหน่งได้อย่างแม่นยำเพื่อความถูกต้องของข้อมูล ซึ่งการระบุตำแหน่งสามารถทำได้หลายวิธี และมีความแม่นยำแตกต่างกันไป

2. วรรณกรรมที่เกี่ยวข้อง

2.1 Global Positioning System

Global Positioning System (GPS) [1] เป็นระบบกำหนดตำแหน่งบนพื้นโลก โดยปัจจัยที่มีผลต่อความแม่นยำในการแสดงพิกัด ได้แก่ ความแปรปรวนของชั้นบรรยากาศ, การบดบังจากกระจก, ละอองฝน, ต้นไม้ใหญ่ เป็นต้น เนื่องจากคลื่นเมื่อตกกระทบกับวัตถุต่างๆ จะเกิดการหักเห ทำให้ได้สัญญาณที่อ่อนลง เหนือสิ่งอื่นใด ประสิทธิภาพของเครื่องรับสัญญาณที่มีความไวในการรับสัญญาณมากน้อยเพียงใด ก็มีผลต่อความแม่นยำด้วย ในงานวิจัยนี้ได้ให้ผู้ทดลองได้ทำการระบุตำแหน่งของตน โดยใช้ค่าพิกัดจาก GPS ในกรณีที่โทรศัพท์เคลื่อนที่ซึ่งทำตัวเป็นเครื่องรับสัญญาณ (Receiver) สามารถรับค่าพิกัดจากดาวเทียมได้ แต่ถ้าไม่สามารถรับค่าพิกัดได้ ก็จะทำให้ผู้ทดลองกำหนดตำแหน่งบนแผนที่เพื่อแสดงพิกัดด้วยตนเอง

2.2 Received Signal Strength Indicator

Received Signal Strength Indicator (RSSI) หรือ ระดับความเข้มของสัญญาณ [2] เป็นตัวชี้วัดความสัมพันธ์กันระหว่างโทรศัพท์เคลื่อนที่ของผู้ใช้งาน กับเสาสัญญาณของผู้ให้บริการ โดย RSSI จะเป็นตัวระบุความเป็นไปได้ในการเชื่อมต่อการสื่อสารไร้สายนี้ ซึ่งขึ้นอยู่กับหลายปัจจัย อาทิเช่น ระยะทางระหว่างโทรศัพท์กับเสาสัญญาณ, ลักษณะภูมิประเทศ, จำนวนผู้ใช้งาน เป็นต้น โดยค่า RSSI จะแปรผันตรงกับความแรงของสัญญาณ มีหน่วยเป็น dBm (decibels milliwatt)

2.3 การระบุตำแหน่งผู้ใช้งาน

มีหลายงานวิจัยที่ใช้ Wireless Technology เพื่อสร้างระบบระบุตำแหน่งของผู้ใช้งาน ซึ่งสามารถแบ่งงานวิจัยตาม Wireless Technology ได้ดังนี้

2.3.1 ใช้ GPS ร่วมกับ Wi-Fi ในการระบุตำแหน่ง

[3] ได้นำเสนอวิธีในการพัฒนาระบบติดตามรถที่ใช้ในงานโลจิสติกส์เพื่อให้มีความแม่นยำในการติดตามมากขึ้น

โดยเฉพาะในเมืองที่มีตึกสูงหรือต้นไม้ใหญ่บดบังสัญญาณ GPS ซึ่งได้ใช้สัญญาณ Fi-Wi ที่ติดตั้งอยู่ตามถนนสายหลักต่างๆเข้ามาช่วยเพิ่มความแม่นยำในการคำนวณ โดยทำการทดลองในเมืองฮ่องกง ที่เต็มไปด้วย Wi-Fi access point สาธารณะที่ติดตั้งตามคูโทรศัพท์และร้านค้าสะดวกซื้อ และจากการที่ได้ข้อมูลตำแหน่งและเวลาของมือถือที่ทำการลองนี้สามารถนำข้อมูลไปหาเส้นทางจราจรแบบเรียลไทม์ต่อไปได้ แต่ทั้งนี้วิธีการนี้มีข้อด้อยคือ เหมาะที่จะใช้กับสถานที่ที่เต็มไปด้วย access point กระจายอยู่รอบๆเส้นทางการเดินทางเท่านั้น

2.3.2 ใช้ Wi-Fi ระบุตำแหน่ง

มิงงานวิจัย [4] และ [5] ที่ได้ใช้ประโยชน์จากสัญญาณ Wi-Fi มาใช้ในการระบุตำแหน่งตัวบุคคล เนื่องจากต้องการที่จะสร้างเครื่องมือแสดงตำแหน่งที่มีประสิทธิภาพจากอุปกรณ์ใกล้ตัว ดังเช่น สมาร์ทโฟน ซึ่งช่วยในการลดค่าใช้จ่าย

งานวิจัย [4] ได้ใช้วิธีในการให้อุปกรณ์มือถือที่ต้องการทราบตำแหน่ง (โหนด A) ทำการรับการ แพร่สัญญาณตำแหน่ง (RSS) ของอุปกรณ์มือถืออีกเครื่อง (โหนด B) ที่มีตำแหน่งที่แน่นอนซึ่งอยู่บริเวณนั้น เพื่อทำการคำนวณตำแหน่งของตน ในขณะเดียวกันเมื่อโหนด A รู้ตำแหน่งที่แน่นอนแล้ว ก็สามารถแพร่สัญญาณค่าตำแหน่งของตนกลับไปยังโหนด B ได้เช่นกัน

ส่วนงานวิจัย [5] ได้สร้างระบบระบุตำแหน่งภายในอาคาร โดยให้มีการสื่อสารกันระหว่าง สมาร์ทโฟนกับ distributed server ซึ่งการทำงานแบ่งเป็นสองส่วน โดยในส่วนแรกให้ผู้ทดลองทำการพิกัดตำแหน่งที่ตนอยู่จาก radiomap ที่ได้รับจากเซิร์ฟเวอร์แล้วเลือกจำนวนของสัญญาณที่จะทำการเก็บค่าในตำแหน่งพิกัดนั้นๆ ซึ่งในส่วนนี้ถือเป็นการนำเทคนิคคราด์ซอร์สซึ่งมาใช้ เมื่อทำการบันทึกค่าสัญญาณเรียบร้อยแล้ว ข้อมูลจะถูกส่งไปเก็บในเซิร์ฟเวอร์ และผู้ใช้งานในส่วนที่สองก็จะสามารถใช้ระบบในการหาตำแหน่งของตนได้จากการใช้ข้อมูลที่มีอยู่ในเซิร์ฟเวอร์ร่วมกับสามารถเลือกอัลกอริทึมเพื่อคำนวณหาตำแหน่งได้ด้วยตนเอง เพราะผู้วิจัยต้องการพิสูจน์ประสิทธิภาพของระบบและอัลกอริทึม นอกจากนี้ค่า RSS ที่ถูกเก็บไว้ในอุปกรณ์มือถือเป็น log files จะถูกส่งกลับไปเก็บไว้ในฐานข้อมูลในเซิร์ฟเวอร์ เพื่อทำการอัปเดต radiomap ต่อไป

แต่ทั้ง 2 งานนี้มีข้อด้อยที่เหมือนกันคือ จำเป็นที่จะต้องมีการใช้อุปกรณ์อีกตัวไม่ว่าจะเป็นสมาร์ทโฟนหรือ Distributed Server

ในการแพร่กระจายสัญญาณ RSS เพื่อให้เครื่องรับสามารถทราบตำแหน่งของตนได้

ทั้งนี้ ผู้วิจัยได้ให้ความสนใจในการพัฒนาระบบเพื่อสร้างแผนที่แสดงค่าความแรงสัญญาณของเครือข่ายต่างๆ ตามตำแหน่งต่างๆ ดังนั้นจึงจำเป็นที่จะต้องมีการระบุตำแหน่งเช่นเดียวกันโดยใช้พิกัดจาก GPS ที่มีความแม่นยำสำหรับภายนอกอาคารเพื่อใช้แสดงตำแหน่งภายนอกอาคาร ร่วมกับการใช้ข้อมูลท้องถิ่น เพื่อแสดงตำแหน่งภายในอาคาร และสามารถเพิ่มข้อมูลตำแหน่งที่ต้องการระบุเพิ่มเติมเองได้ ซึ่งข้อดีคือสามารถเพิ่มระบุตำแหน่งได้เอง และไม่จำเป็นที่จะต้องมียุทธศาสตร์ของ Wi-Fi access point เพื่อใช้ในการคำนวณซึ่งทำให้สามารถทำการเก็บข้อมูลที่ใดก็ได้

2.4 Crowdsourcing

การสำรวจสัญญาณของเครือข่ายต่างๆ ในงานวิจัยนี้ ได้ใช้เทคนิค คราวด์ซอร์ซซิ่ง (Crowdsourcing) หรือการใช้พลังมวลชน ซึ่งเป็นการกระจายงานหรือปัญหาให้คนจำนวนมากช่วยกันทำงานหรือแก้ปัญหา ตัวอย่างเช่นการช่วยกันรายงานสภาพจราจรไปยังสถานีวิทยุก็ถือว่าเป็นคราวด์ซอร์ซซิ่งแบบหนึ่ง ประโยชน์จากการใช้เทคนิคคราวด์ซอร์ซซิ่งทำให้ได้คำตอบที่หลากหลายและช่วยลดเวลาและค่าใช้จ่ายในการสำรวจ ปัจจุบันการใช้อินเทอร์เน็ตทำให้สามารถใช้เทคนิคคราวด์ซอร์ซซิ่งได้สะดวกและเข้าถึงผู้คนจำนวนมาก ดังตัวอย่างเช่น โครงการ Read for the Blind [6] ที่เปิดโอกาสให้ทุกคนร่วมกันสร้างหนังสือเสียงผ่านแอปพลิเคชันเพื่อคนตาบอดหรือ การแข่งขัน Test Bug Battle ที่ให้ผู้ทดสอบจาก 60 กว่าประเทศค้นหา Bugs จาก Web Browsers [7] เป็นต้น

2.5 Coverage Map

Coverage map เป็นการแสดงพื้นที่ที่สัญญาณเครือข่ายต่างๆ สามารถให้บริการได้ โดยงานวิจัยนี้ได้แสดงผลสัญญาณของเครือข่ายที่สามารถรับได้ ณ บริเวณนั้นๆ ในรูปแบบ Heat map ซึ่งเป็นการแสดงข้อมูลโดยใช้การลำดับสีในการนำเสนอค่าของข้อมูลในรูปแบบภาพ 2 มิติ ซึ่งเป็นเทคนิคการแสดงผลข้อมูลอีกรูปแบบหนึ่ง แทนการแสดงข้อมูลแบบชาร์ตและแบบกราฟ [8], [9]

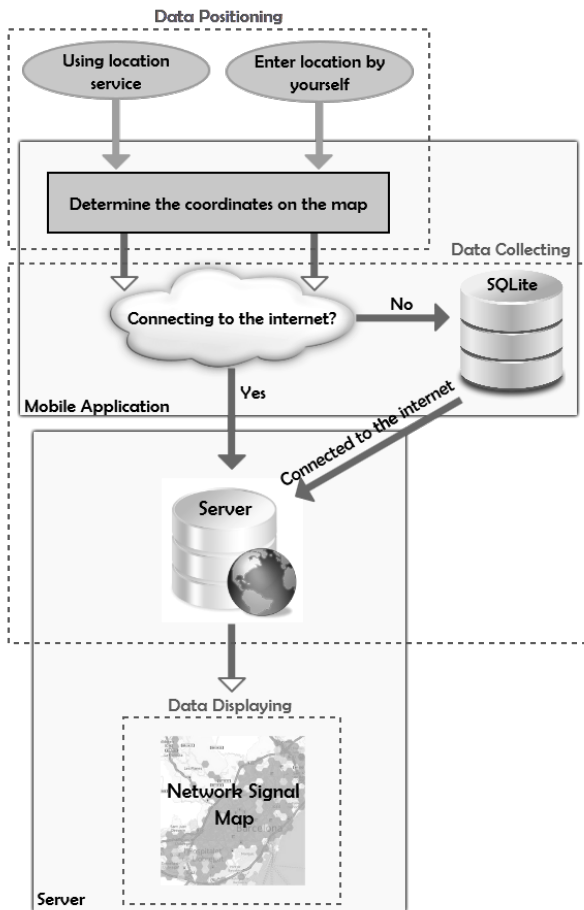
ในปัจจุบันได้มีงานที่ทำการแสดงผลแผนที่ของสัญญาณเครือข่ายในรูปแบบต่างๆ เช่น Voronoi diagram [10], Grid map

[11], Heat map [12] เป็นต้น โดยงานที่ผ่านมาซึ่งทำการแสดงผลแผนที่ของสัญญาณในรูปแบบ Heat map นั้น ได้แก่ OpenSignal [13], Sensorly [14], Skyhook [15] เป็นต้น โดยทั้ง 3 งานนี้ได้ใช้เทคนิคคราวด์ซอร์ซซิ่งในการช่วยเก็บข้อมูล โดยได้เผยแพร่ระบบให้ผู้ใช้ช่วยกันทำการเก็บข้อมูล และสามารถทำได้ข้อมูลมาจากทั่วโลก ทั้งข้อมูลความแรงสัญญาณโทรศัพท์เคลื่อนที่ และข้อมูล Wi-Fi แต่ทั้งนี้ทั้ง 3 ระบบก็มีข้อด้อยคือแผนที่ที่แสดงผลข้อมูลของทั้ง 3 งานนั้นไม่สามารถแสดงได้ในระดับที่ละเอียดเพียงพอหรือในระดับห้อง นอกจากนี้ไม่สามารถแสดงผลข้อมูลท้องถิ่นได้ เช่น ค่าความแรงสัญญาณจาก Wi-Fi access point ของแต่ละพื้นที่ เป็นต้น และไม่สามารถแสดงได้ถึงความพร้อมใช้งานของเครือข่าย (Network Availability) และความสามารถในการเชื่อมต่อเครือข่าย (Network Connectivity) ณ บริเวณนั้นๆ

ทั้งนี้ในงานวิจัยนี้ได้สร้างแผนที่แสดงความน่าจะเป็นของความแรงสัญญาณทั้งในรูปแบบ Network Availability และ Network Connectivity ของ Wi-Fi และ โทรศัพท์เคลื่อนที่ ซึ่งสามารถให้ความละเอียดที่เพียงพอทั้งภายนอกอาคารและภายในอาคาร รวมถึง Wi-Fi ที่ทำการเก็บข้อมูลมานั้น เป็น Wi-Fi ระดับท้องถิ่น โดยได้ใช้เทคนิคคราวด์ซอร์ซซิ่งในการช่วยเก็บข้อมูลและแสดงผล ผ่านโมบายล์แอปพลิเคชันบนระบบแอนดรอยด์ และสามารถทำการเก็บข้อมูลได้ทั้งในสถานะออนไลน์และออฟไลน์ ซึ่งจะทำให้สามารถเก็บข้อมูลแม้ในที่อับสัญญาณ หรือในขณะที่ไม่มีการเชื่อมต่ออินเทอร์เน็ตตามความต้องการของผู้ใช้งานได้

3. ขั้นตอนวิธีการดำเนินการวิจัย

ในงานวิจัยนี้ได้แบ่งขั้นตอนการทำงานของระบบออกเป็น 3 ส่วน ดังนี้



ภาพที่ 1: ขั้นตอนการทำงานของระบบ

3.1 การแสดงพิกัดตำแหน่ง (Data Positioning)

เป็นส่วนของการระบุตำแหน่งบริเวณที่จะทำการสำรวจสัญญาณเครือข่าย และเนื่องจากกรณีที่เครื่องมือจะต้องสามารถทำการสำรวจได้ทั้งภายในอาคารและนอกอาคาร จึงได้ทำการออกแบบโดยแบ่งส่วนแสดงพิกัดตำแหน่งเป็น 2 ส่วน คือ

3.1.1 ส่วนที่รับค่าพิกัดจาก Location Service

เป็นส่วนที่จะถูกเปิดใช้เมื่อแอปพลิเคชันสามารถรับค่าพิกัดจากดาวเทียมได้แล้วจะทำการรับค่าพิกัดนี้ไปเรื่อยๆ ตลอดเวลาที่เปิดใช้งาน (tracking)

3.1.2 ส่วนที่รับค่าจากผู้ใช้งานเอง

เป็นส่วนที่จะถูกเปิดใช้เมื่อแอปพลิเคชันไม่สามารถรับค่าพิกัดจากดาวเทียมได้ โดยการเก็บข้อมูลแต่ละครั้งในแต่ละตำแหน่ง ผู้ทดลองจะต้องทำการระบุตำแหน่งด้วยตัวเอง

โดยการระบุตำแหน่งด้วยทั้ง 2 วิธีดังที่กล่าวมานี้ จะทำการแสดงบนแผนที่ Google Maps

3.2 การเก็บข้อมูล (Data Collecting)

การเก็บข้อมูลเป็นส่วนที่จะทำงานไปพร้อมกับในส่วนแสดงพิกัดตำแหน่ง (Data Positioning) ซึ่งทำการเก็บข้อมูลต่างๆของสัญญาณเครือข่ายที่เครื่องมือสามารถมองเห็นได้ ณ ขณะนั้น ในกรณีที่เครื่องมือสามารถรับสัญญาณได้ทั้ง Wi-Fi และสัญญาณจากโทรศัพท์เคลื่อนที่พร้อมกัน ก็จะสามารถเก็บบันทึกข้อมูลของสัญญาณได้ทั้ง 2 ประเภท

หลังจากที่ได้ข้อมูลมาแล้ว ระบบก็จะทำการเลือกแหล่งเก็บข้อมูลซึ่งเป็นการทำงานแบบไดนามิก โดยที่จะทำการตรวจสอบก่อนว่าในขณะนั้น สมาร์ทโฟนมีการเชื่อมต่ออินเทอร์เน็ตอยู่หรือไม่ ถ้ามีการเชื่อมต่ออยู่ก็จะทำการส่งข้อมูลที่ได้ไปยังเซิร์ฟเวอร์ เพื่อทำการเก็บข้อมูลเข้าฐานข้อมูล MySQL ต่อไป แต่ถ้าไม่มีการเชื่อมต่ออินเทอร์เน็ต ข้อมูลก็จะถูกนำไปเก็บไว้ใน SQLite ก่อน แล้วเมื่อใดก็ตามที่มีการเชื่อมต่ออินเทอร์เน็ต เครื่องมือก็จะทำการส่งข้อมูลที่เก็บใน SQLite นี้ออกไปเก็บไว้ใน MySQL บนเซิร์ฟเวอร์แล้วทำการลบข้อมูลใน SQLite ออก

3.3 การแสดงผลข้อมูล (Data Display)

เป็นส่วนที่จะนำข้อมูลที่บันทึกในฐานข้อมูลมาประมวลผลเพื่อแสดงเป็นข้อมูลความน่าจะเป็นของความแรงสัญญาณ โดยมีการคำนวณข้อมูลเพื่อนำไปแสดงผล 2 รูปแบบ ดังนี้

3.3.1 Network Availability

ในงานวิจัยนี้ได้ให้ความสำคัญกับนิยาม Network Availability ไว้ว่า คือความน่าจะเป็นที่อุปกรณ์สื่อสารสามารถรับสัญญาณของเครือข่ายต่างๆได้ ณ บริเวณใดบริเวณหนึ่ง โดยสัญญาณ Wi-Fi และสัญญาณโทรศัพท์เคลื่อนที่ ที่อุปกรณ์สามารถรับได้นั้นต้องไม่น้อยกว่าค่าคงที่ค่าหนึ่ง ซึ่งจากนิยามสามารถเขียนเป็นสมการได้ ดังนี้

$$\text{Network Availability} = \frac{\sum_{i=1}^n (a_1 + a_2 + \dots + a_n)}{n} \quad (1)$$

โดยที่ n คือ จำนวนครั้งที่ขึ้น ณ ตำแหน่งๆหนึ่ง, i คือ ลำดับครั้งที่ขึ้น ณ ตำแหน่งนั้น, a คือ จำนวนเครือข่ายที่อุปกรณ์สามารถรับได้ในแต่ละครั้งที่ขึ้น ณ ตำแหน่งนั้น โดยสัญญาณ Wi-Fi และสัญญาณโทรศัพท์เคลื่อนที่ ที่อุปกรณ์สามารถรับได้นั้นต้องไม่น้อยกว่าค่าคงที่ C

3.3.2 Network Connectivity

ในงานวิจัยนี้ ได้ให้ความหมายของ Network Connectivity ไว้ว่า คือ ความน่าจะเป็นที่อุปกรณ์สื่อสารสามารถเชื่อมต่อกับเครือข่ายสัญญาณใดๆ ได้ ณ บริเวณใดบริเวณหนึ่ง โดยจะพิจารณาคำนวณเครือข่ายเป็นรายๆ ไป ซึ่งจากนิยามสามารถเขียนเป็นสมการได้ ดังนี้

$$\text{Network Connectivity} = \frac{\sum_{i=1}^n (c_1 + c_2 + \dots + c_n)}{n} \quad (2)$$

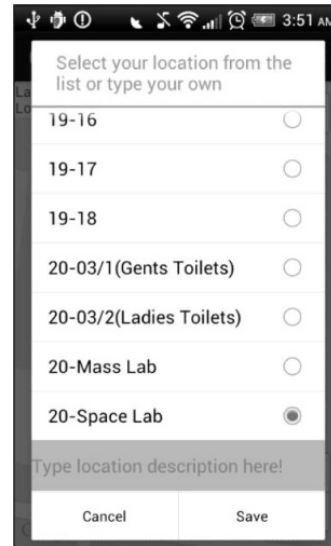
โดยที่ n คือ จำนวนครั้งที่เก็บข้อมูล ณ ตำแหน่งๆ หนึ่ง, i คือ ลำดับครั้งที่เก็บข้อมูล ณ ตำแหน่งนั้น, c จะมีค่าเป็น 1 เมื่อสามารถทำการเชื่อมต่อกับเครือข่ายนั้นได้ และจะมีค่าเป็น 0 เมื่อไม่สามารถเชื่อมต่อเครือข่ายได้

4. ผลการดำเนินการวิจัย

เพื่อให้ได้มาซึ่งแผนที่สัญญาณเครือข่ายต่างๆ จึงได้สร้างแอปพลิเคชันบนระบบแอนดรอยด์ขึ้นเพื่อทำการทดลองสำรวจสัญญาณของเครือข่าย โดยหน้าจอของระบบเป็นดังภาพที่ 2 – 3 และหน้าผลการทดลองเป็นดังภาพที่ 4 - 5 โดยภาพที่ 4 เป็นการแสดงผลจากค่าความแรงสัญญาณ ณ บริเวณที่ทำการสำรวจในรูปแบบ Heat map และภาพที่ 5 เป็นการแสดงผลความน่าจะเป็นของความแรงสัญญาณ ณ บริเวณนั้น ในรูปแบบ Grid heat map



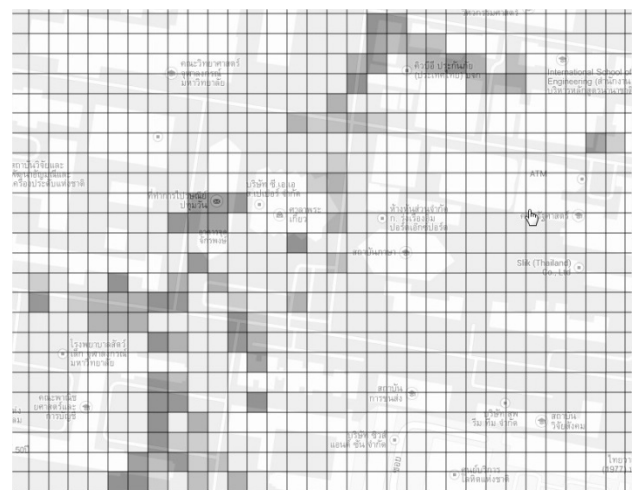
ภาพที่ 2: หน้าจอของระบบที่ใช้สำรวจรับค่าพิกัดจาก GPS



ภาพที่ 3: หน้าจอของระบบที่ใช้สำรวจเลือกตำแหน่งด้วยตนเอง



ภาพที่ 4: หน้าจอผลการสำรวจในรูปแบบ Heat map



ภาพที่ 5: หน้าจอผลการสำรวจในรูปแบบ Grid heat map

5. สรุป

ผู้วิจัยได้พัฒนาแอปพลิเคชันบนระบบแอนดรอยด์ขึ้นเพื่อทำการสำรวจข้อมูลสัญญาณของเครือข่ายต่างๆ โดยสามารถทำการเก็บข้อมูลได้ทั้งภายนอกและภายในอาคาร เนื่องจากใช้การระบุตำแหน่งทั้งแบบ GPS และการระบุตำแหน่งด้วยตนเองในการเก็บข้อมูล เพื่อที่จะสามารถเก็บข้อมูลได้ในทุกตำแหน่งตามที่ผู้ใช้งานต้องการ และสามารถเก็บข้อมูลได้ทั้งในขณะที่ออนไลน์และออฟไลน์

หลังจากที่ทำการสำรวจแล้ว สามารถนำข้อมูลนี้มาแสดงเพื่อที่ต้องการให้เป็นตัวช่วยให้ผู้ใช้งานได้ทราบว่ามีบริเวณใดหรือเครือข่ายใดมีการใช้งานของสัญญาณที่ดี เพื่อประกอบการตัดสินใจ และเพื่อที่ผู้ให้บริการจะนำไปปรับปรุง แก้ไขสัญญาณให้ดีขึ้นต่อไป

ในส่วนองงานในอนาคต ควรจะมีการปรับปรุงพัฒนาระบบที่ลดการใช้พลังงานแบตเตอรี่ของสมาร์ตโฟน เพื่อที่ผู้ใช้งานจะสามารถเปิดแอปพลิเคชันเก็บข้อมูลได้ตลอดเวลา

เอกสารอ้างอิง

- [1] GPS.GOV, "Systems," [online] 2013.
Availablefrom: <http://www.gps.gov/systems/gps/> [2013,September 10].
- [2] Wikipedia, "Heat map," [online] 2013.
Availablefrom: http://en.wikipedia.org/wiki/Received_signal_strength_indication [2013, September 10].
- [3] Mok, E. C. M., "Using outdoor public WiFi and GPS integrated method for position updating of knowledge-based logistics system in dense high rise urban environments." *Supply Chain Management and Information Systems*, 2010.
- [4] Pourabdollah, A., M. Xiaolin, et al., "Towards low-cost collaborative mobile positioning." *Ubiquitous Positioning Indoor Navigation and Location Based Service*, 2010.
- [5] Laoudias, C., G. Constantinou, et al., "The Airplace Indoor Positioning Platform for Android Smartphones." *Mobile Data Management*, 2012.
- [6] Facebook, "Read for the Blind," [online] 2013.
Availablefrom: <https://www.facebook.com/readfortheblind> [2014, March 15].
- [7] uTest Inc., "Software Testing Blog," [online] 2008.
Availablefrom: <http://blog.utest.com/bug-battle-results/2008/12/> [2014, March 15].
- [8] Wikipedia, "Heat map," [online] 2013.
Availablefrom: http://en.wikipedia.org/wiki/Heat_map [2013, November 11].
- [9] Lab Escape Inc., "What Is A Heat Map?," [online] 2004.
Availablefrom: <http://www.labescape.com/info/articles/what-is-a-heat-map.html> [2013, November 11].
- [10] Park, J.-g., Charrow, B., Curtis, D., Battat, J., Minkov, E., Hicks, J., Teller, S., and Ledlie, J., "Growing an organic indoor location system," *In Proc. Intl. Conf. on Mobile Systems, Applications, and Services*, pp. 271–284, 2010.
- [11] Y. Kim, Y. Chon, and H. Cha, "Smartphone-Based Collaborative and Autonomous Radio Fingerprinting," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, pp. 112-122, 2012.
- [12] Y. S. Jeong, Y. J. Chung, and J. H. Park, "Visualisation of efficiency coverage and energy consumption of sensors in wireless sensor networks using heat map," *IET Commun.*, pp. 1129–1137, May 2011.
- [13] OpenSignal, "OpenSignal," [online] 2013.
Availablefrom: <http://opensignal.com/> [2013, July 2].
- [14] Sensorly, "Sensorly" [online] 2013.
Availablefrom: <http://www.sensorly.com/> [2013, July 2].
- [15] Skyhook Wireless, "Skyhook" [online] 2013.
Availablefrom: <http://www.skyhookwireless.com/> [2013, July 2].

ระบบแจ้งเตือนอุบัติเหตุและการระบุตำแหน่งบนแผนที่โดยการใช้อุปกรณ์เคลื่อนที่ สำหรับบริการทางการแพทย์ฉุกเฉิน

Accident Warning System and Specifying the Location on a Map using Mobile Device for the Emergency Medical Services

ทวิชชัย เรืองธนาบุรีรักษ์ (Tawatchai Ruangtanurk)¹ งามนิจ อาจอินทร์ (Ngamnij Arch-int)²

สมจิตร อาจอินทร์ (Somjit Arch-int)³ สายย์ณู สายยศ (Saiyan Saiyod)⁴ และปรีวัฒน์ ภู่งเงิน (Pariwat Phu-ngern)⁵

^{1,2,3,4}ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น

⁵ภาควิชาเวชศาสตร์ฉุกเฉิน คณะแพทยศาสตร์ มหาวิทยาลัยขอนแก่น

¹tawatchai_ruang@kkumail.com, ²ngamnij@kku.ac.th, ³somjit@kku.ac.th, ⁴saiyan@kku.ac.th, ⁵pariwat555@hotmail.com

บทคัดย่อ

ภาวะฉุกเฉินที่เกิดขึ้นบ่อยครั้งที่สุดและมีสถิติที่เป็นสาเหตุในการเสียชีวิตของคนไทยสูงมากได้แก่ อุบัติเหตุ ซึ่งปกติการช่วยเหลือผู้ประสบอุบัติเหตุบนท้องถนนโดยหน่วยกู้ภัยหรือรถฉุกเฉินของโรงพยาบาลจำเป็นต้องไปถึงยังจุดเกิดเหตุให้เร็วที่สุดเพื่อนำผู้ป่วยไปส่งยังโรงพยาบาลที่อยู่ใกล้ที่สุดได้ ซึ่งในปัจจุบันจะพบปัญหาในการแจ้งเหตุฉุกเฉิน เช่นการแจ้งสถานที่เกิดเหตุที่คลาดเคลื่อน การไม่ชำนาญทางของผู้ขับรถฉุกเฉิน ทำให้ไม่สามารถช่วยเหลือผู้ป่วยได้ทันท่วงที งานวิจัยนี้จึงได้นำเสนอการออกแบบและพัฒนาระบบแจ้งเตือนอุบัติเหตุและการระบุตำแหน่งบนแผนที่สำหรับบริการทางการแพทย์ฉุกเฉิน โดยจะช่วยให้ผู้พบเห็นอุบัติเหตุหรือตัวผู้ประสบอุบัติเหตุเองสามารถเรียกใช้ Application บนโทรศัพท์มือถือ (smart phone) เพื่อส่งข้อมูลสถานที่เกิดเหตุ เช่นพิกัดหรือตำแหน่งของผู้ประสบอุบัติเหตุ รวมไปถึงข้อมูลอื่นๆที่เกี่ยวข้องไปยังศูนย์กลางรับแจ้งเหตุประจำจังหวัดและโรงพยาบาลที่อยู่ในพื้นที่เกิดเหตุโดยอัตโนมัติ เพื่อส่งข้อมูลดังกล่าวต่อไปยังรถรับส่งผู้ป่วยฉุกเฉินของโรงพยาบาลที่ว่างและพร้อมปฏิบัติหน้าที่ และให้รถฉุกเฉินของโรงพยาบาลทราบตำแหน่งจุดเกิดเหตุในลักษณะของแผนที่ ได้อย่างชัดเจนและแม่นยำ ทั้งมีการแนะนำเส้นทางที่ทำให้รถฉุกเฉินของโรงพยาบาลสามารถไปรับผู้ประสบอุบัติเหตุได้อย่างรวดเร็วและทันการณ์ รวมไปถึงการบอกเส้นทางไปยังโรงพยาบาลที่ใกล้ที่สุดได้

คำสำคัญ: การแพทย์ฉุกเฉิน เวชศาสตร์ฉุกเฉิน ระบบแจ้งเตือนอุบัติเหตุ การระบุตำแหน่งบนแผนที่

Abstract

The emergency that occurs most frequently and ranks as the highest cause of death among Thai people is traffic accidents. Normally, to rescue the victims on roads, the Salvation Army or an ambulance is required to reach the scene as quickly as possible in order to transport the victims to the nearest hospital. However, some problems regarding unclear information, such as incorrect location or unfamiliarity on the part of the ambulance driver with the location, may cause a delay. This research is intended to present a system designed to provide the correct information and location to the emergency personnel. An eyewitness or the victim will be able to use an application on his or her smart phone to automatically inform the local provincial or hospital call center of the location of the accident along with any other related information. This information will alert an available ambulance to access the exact location by using a online map. The route to the accident's location is also provided in order that the ambulance driver can quickly drive to the site. Additionally, the route to the nearest hospital is also provided.

Keyword: Emergency Medical, Accidental Warning System, Global Positioning System.

1. บทนำ

การแพทย์ฉุกเฉินหรือเวชศาสตร์ฉุกเฉินเป็นการแพทย์เฉพาะทางที่มุ่งเน้นการวินิจฉัยและรักษาความเจ็บป่วยหรือการบาดเจ็บที่เกิดขึ้นฉับพลันและต้องการความช่วยเหลือทางการแพทย์อย่างเร่งด่วน การแจ้งเตือนเพื่อทำการช่วยเหลือผู้ป่วยได้ทันเวลาที่จึงเป็นปัจจัยสำคัญต่อชีวิตของผู้ประสบอุบัติเหตุ ซึ่งปกติการช่วยเหลือเหตุฉุกเฉินของประเทศไทยในปัจจุบันนั้นเริ่มจากการรับแจ้งเหตุฉุกเฉินเพื่อขอความช่วยเหลือด้วยระบบบริการทางการแพทย์ฉุกเฉินผ่านทางโทรศัพท์หมายเลข 1669 โดยจะติดต่อไปยังศูนย์สื่อสารซึ่งเป็นศูนย์สั่งการที่ตั้งอยู่จังหวัดละ 1 แห่ง โดยเจ้าหน้าที่จะสอบถามถึงข้อมูลและรายละเอียดที่สำคัญต่อการเข้าช่วยเหลือเหตุฉุกเฉิน เช่น สถานที่เกิดเหตุ, ข้อมูลผู้ป่วย และอาการเบื้องต้นของผู้ป่วย จากนั้นเจ้าหน้าที่จะพิจารณาว่าเหตุฉุกเฉินดังกล่าวเกิดขึ้นภายในพื้นที่รับผิดชอบของโรงพยาบาลใด จากนั้นจึงใช้วิทยุสื่อสารเพื่อแจ้งข้อมูลเหตุฉุกเฉินดังกล่าวไปยังโรงพยาบาลนั้น (หรือในบางกรณีจะมีการโทรแจ้งเหตุฉุกเฉินมายังศูนย์รับแจ้งเหตุฉุกเฉินของโรงพยาบาลโดยตรง) ซึ่งในปัจจุบันจะพบปัญหาในการรับแจ้งเหตุฉุกเฉินและการช่วยเหลือผู้ประสบอุบัติเหตุที่สามารถแบ่งได้เป็นประเด็นต่างๆดังต่อไปนี้

1.1 การช่วยเหลือผู้ป่วยจากเหตุฉุกเฉินโดยระบบเดิม จะใช้เพียงวิทยุสื่อสารเป็นเครื่องมือในการแจ้งเหตุและประสานงาน ซึ่งจะมีความล่าช้าและมีโอกาสที่จะประสบปัญหาขัดข้องของระบบการสื่อสารดังกล่าว ทำให้การติดต่อสื่อสารเพื่อประสานงานทำได้ไม่เต็มที่

1.2 ในการแจ้งข้อมูลเหตุฉุกเฉินและระบุสถานที่เกิดเหตุด้วยระบบเดิม อาจมีการทำงานที่ซ้ำซ้อนหรือจากอุปสรรคในการสื่อสารทำให้ได้รับการแจ้งสถานที่เกิดเหตุคลาดเคลื่อนไม่ชัดเจน กระทั่งเกิดความเข้าใจที่ไม่ตรงกัน จนบางครั้งหน่วยกู้ภัยที่เข้าไปช่วยเหลือผู้ประสบอุบัติเหตุเดินทางไปผิดที่หรือคนขับรถอาจไม่ทราบทางจึงทำให้การช่วยเหลือเป็นไปด้วยความล่าช้า

1.3 ในการรับผู้ประสบอุบัติเหตุไปส่งโรงพยาบาลในเขตพื้นที่ที่ประสบอุบัติเหตุ บางครั้งถ้าผู้ขับรถไม่ชำนาญทางในพื้นที่ที่เกิดเหตุ ก็จะเป็นการยากที่จะรู้ว่าจากจุดเกิดเหตุไปยัง

โรงพยาบาลจะต้องใช้เส้นทางใดที่สั้นและดีที่สุด ที่จะสามารถนำผู้ประสบอุบัติเหตุไปส่งยังโรงพยาบาลได้อย่างรวดเร็ว

จากปัญหาดังกล่าวงานวิจัยนี้จึงได้นำเสนอการออกแบบและพัฒนาระบบแจ้งเตือนอุบัติเหตุและการระบุตำแหน่งบนแผนที่สำหรับบริการทางการแพทย์ฉุกเฉิน โดยจะช่วยให้เจ้าหน้าที่ผู้รับแจ้งเหตุฉุกเฉินสามารถค้นหาหรือระบุตำแหน่งสถานที่เกิดเหตุบนแผนที่ได้อย่างชัดเจน รวดเร็ว พร้อมทั้งทราบว่าเหตุฉุกเฉินเกิดขึ้นบนพื้นที่รับผิดชอบของโรงพยาบาลใด เพื่อส่งข้อมูลและตำแหน่งของจุดเกิดเหตุให้กับศูนย์รับแจ้งเหตุของโรงพยาบาลนั้น นอกจากนี้ยังสามารถจัดการรถฉุกเฉินของโรงพยาบาลเพื่อค้นหารถฉุกเฉินที่พร้อมและสามารถเข้าช่วยเหลือเหตุฉุกเฉินได้รวดเร็วที่สุด รวมถึงการส่งข้อมูลสำคัญของผู้ป่วยไปยังรถฉุกเฉินของโรงพยาบาลคันดังกล่าว ซึ่งภายในรถฉุกเฉินของโรงพยาบาลจะมีการติดตั้งแอปพลิเคชัน (Application) ที่ทำงานบนแท็บเล็ต (Tablet) ซึ่งเป็นอุปกรณ์เคลื่อนที่ เพื่อแสดงจุดเกิดเหตุในลักษณะของแผนที่ ทำให้ทราบตำแหน่งของจุดเกิดเหตุได้อย่างชัดเจนและแม่นยำ ทั้งยังมีระบบการนำทางที่ช่วยให้รถฉุกเฉินของโรงพยาบาลสามารถไปรับผู้ประสบอุบัติเหตุได้อย่างรวดเร็วและทันการณ์ นอกจากนี้ เมื่อรถฉุกเฉินของโรงพยาบาลไปพบผู้ประสบอุบัติเหตุแล้วจะแสดงตำแหน่งของโรงพยาบาลและแผนที่ที่สามารถนำผู้ป่วยส่งไปรักษาได้อย่างรวดเร็ว ซึ่งตลอดการทำงานของระบบนั้น ทางศูนย์รับแจ้งเหตุของโรงพยาบาลสามารถติดตามตำแหน่งของรถฉุกเฉินและติดตามความคืบหน้าของการช่วยเหลือเหตุฉุกเฉินได้อย่างตลอดเวลา

งานวิจัยนี้จะนำเทคโนโลยีเว็บเซอร์วิส (Web Service) [1] มาใช้ โดยจะใช้ภาษาเอ็กซ์เอ็มแอล (Extensible Markup Language: XML) [2] เข้ามาช่วยในการเก็บข้อมูลชั่วคราวและแลกเปลี่ยนข้อมูลระหว่างแพลตฟอร์ม (Platform) ที่แตกต่างกัน และใช้แมปส์ เอพีไอ (Maps API) ในการสร้างและใช้งานแผนที่ รวมถึงระบบกำหนดตำแหน่งบนโลก (Global Positioning System) เพื่อการระบุตำแหน่งของรถฉุกเฉินของโรงพยาบาล

2. ระเบียบวิธีวิจัย

2.1 ศึกษาทฤษฎีที่เกี่ยวข้อง

2.1.1 เว็บเซอร์วิส (Web Services)

เว็บเซอร์วิส คือ แอปพลิเคชัน (Application) หรือ โปรแกรม (Program) ที่ทำงานอย่างใดอย่างหนึ่ง ในลักษณะให้บริการ โดยจะถูกเรียกใช้งานจากแอปพลิเคชันอื่นๆ ในรูปแบบ อาร์ พี ซี (Remote Procedure Call: RPC) ซึ่งการให้บริการจะมีเอกสารที่อธิบายคุณสมบัติของบริการกำกับไว้ โดยภาษาที่ถูกใช้เป็นตัวกลางในการแลกเปลี่ยน คือ เอ็กซ์ เอ็ม แอล ที่อยู่ในรูปแบบของ โพรโตคอล SOAP ที่ถูกใช้เป็นมาตรฐานในการแลกเปลี่ยนข้อมูลของเว็บเซอร์วิส

2.1.2 แมปส์ เอพีไอ (Maps API) [3]

เอพีไอ (Application Programming Interface: API) เป็นช่องทางในการเชื่อมต่อกับเว็บไซต์ผู้ให้บริการเอพีไอจากที่อื่น ช่วยให้ระบบที่อยู่ต่างระบบกันสามารถติดต่อกันได้ โดยจะทำหน้าที่เป็นตัวกลางให้โปรแกรมประยุกต์เชื่อมต่อกับ โปรแกรมประยุกต์อื่นๆ เอพีไอเป็นกลุ่มของฟังก์ชัน (function) หรือ คลาส (Class) ที่ผู้ให้บริการสร้างขึ้น เพื่อรองรับการเรียกข้อมูลจากโปรแกรม

แมปส์ เอพีไอ เป็นการให้บริการ เอพีไอในอีกรูปแบบหนึ่งที่สามารถนำข้อมูลแผนที่จาก เอพีไอที่ผู้ให้บริการได้เปิดไว้มาใช้ประโยชน์ โดยส่วนมากจะใช้กับเว็บไซต์ห้างร้านต่างๆ เพื่อเป็นช่องทางให้ลูกค้าทราบถึงที่อยู่ทางภูมิศาสตร์ของร้าน แต่มากกว่านั้น ปัจจุบัน แมปส์ เอพีไอ มีความสามารถมากมายสามารถทำงานบนแผนที่ทางภูมิศาสตร์ได้อย่างครอบคลุม

2.2 ศึกษางานวิจัยที่เกี่ยวข้อง

2.2.1 Real-time Ambulance Location Monitoring using GPS and Maps Open API [5]

งานวิจัยนี้ได้นำเสนอ ระบบตรวจสอบตำแหน่งที่ตั้งของรถพยาบาลแบบเรียลไทม์ โดยการติดตามตำแหน่งของรถพยาบาลในระบบดังกล่าวจะเชื่อมต่อกันผ่านทางอินเทอร์เน็ตไร้สาย ซึ่งระบบบนรถพยาบาลฉุกเฉินและระบบตรวจสอบบนศูนย์สั่งการจะใช้ข้อมูลแผนที่ของ กูเกิล แมปส์ จากส่วนเซิร์ฟเวอร์แผนที่ ซึ่งระบบรถพยาบาลฉุกเฉินจะได้รับข้อมูลทางภูมิศาสตร์จากระบบกำหนดตำแหน่งบนโลก และจะส่งพิกัดตำแหน่งไปที่ระบบตรวจสอบบนศูนย์สั่งการจากการ

สื่อสารผ่านเครือข่ายไร้สาย จากนั้นระบบตรวจสอบบนศูนย์สั่งการจะแสดงที่ตั้งของรถพยาบาลบนแผนที่

2.2.2 Mobile Emergency System and integration [6]

งานวิจัยนี้ได้นำเสนอระบบฉุกเฉินทางการแพทย์ขั้นสูง เพื่ออำนวยความสะดวกในการทำงานและการประมวลผลของกระบวนการทั้งหมดที่เกี่ยวข้องกับการให้บริการในกรณีฉุกเฉิน จากการแจ้งการเกิดเหตุผ่านแอปพลิเคชันบนมือถือและส่งข้อมูลไปยังรถฉุกเฉินของโรงพยาบาลที่พร้อมเข้าช่วยเหลือเพื่อบอกตำแหน่งของเหตุฉุกเฉิน และทำการค้นหาโรงพยาบาลที่ใกล้ที่สุด ซึ่งงานวิจัยชิ้นนี้เป็นแนวคิดที่กำลังได้รับการสนใจ และเป็นการออกแบบที่อยู่ในช่วงของการพัฒนา

2.3 วิธีการดำเนินงาน

2.3.1 การวิเคราะห์ความต้องการ (Requirement Analysis)

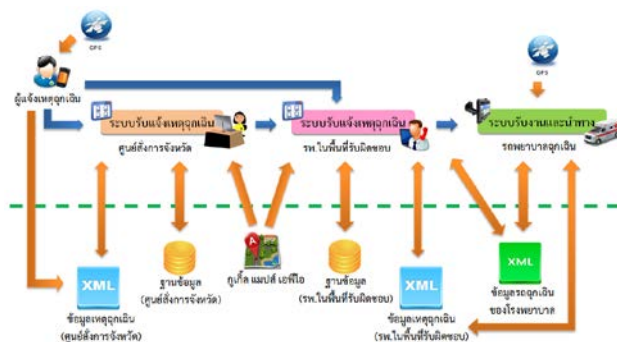
เป็นการศึกษารายละเอียดและขั้นตอนการทำงานของระบบบริการการแพทย์ฉุกเฉิน (Emergency medical services) ของประเทศไทยในปัจจุบัน จากองค์ประกอบ คือ กลไกการสื่อสาร เพื่อตอบสนองต่อความต้องการทางการแพทย์ฉุกเฉิน (Communications mechanism to initiate a response), พาหนะและบุคลากรผู้ให้บริการ (Vehicle with personnel) เพื่อให้การรักษาและนำส่งผู้ป่วย รวมถึงสถานพยาบาลที่รองรับผู้ป่วย (Receiving facility) เพื่อเป็นข้อมูลในการวิเคราะห์ปัญหาและความต้องการของระบบงานบริการการแพทย์ฉุกเฉิน รวมไปถึงวิเคราะห์องค์ประกอบของฮาร์ดแวร์และซอฟต์แวร์ที่สนับสนุน

2.3.2 การออกแบบระบบ (System Design)

ในการออกแบบสถาปัตยกรรมของระบบแจ้งเตือนอุบัติเหตุและการระบุตำแหน่งบนแผนที่โดยการใช้อุปกรณ์เคลื่อนที่สำหรับบริการทางการแพทย์ฉุกเฉิน เพื่อแสดงองค์ประกอบที่เป็นภาพรวมทั้งหมดของระบบ การรับส่งข้อมูลและความสัมพันธ์กันระหว่างองค์ประกอบต่างๆ ของสถาปัตยกรรม ดังแสดงในภาพที่ 1 ซึ่งจะประกอบด้วยส่วนของแหล่งเก็บข้อมูลและส่วนติดต่อกับผู้ใช้งาน

ระบบการรับแจ้งเหตุฉุกเฉินเป็นแอปพลิเคชันที่ทำงานอยู่ ณ ศูนย์สั่งการจังหวัด และศูนย์รับแจ้งเหตุฉุกเฉินของโรงพยาบาล โดยมีเจ้าหน้าที่ประจำศูนย์รับแจ้งเหตุฉุกเฉินเป็น

ผู้ใช้งาน การแจ้งเหตุฉุกเฉินจะเกิดขึ้นโดยผู้แจ้งเหตุฉุกเฉินใช้สมาร์ตโฟนที่เชื่อมต่อกับเครือข่ายอินเทอร์เน็ตและโมดูลของระบบกำหนดตำแหน่งบนโลก โดยจะต้องกรอกข้อมูลที่จำเป็นต่อการเข้าช่วยเหลือเหตุฉุกเฉินในแอปพลิเคชัน ซึ่งแอปพลิเคชันจะทำการดึงข้อมูลตำแหน่งปัจจุบันของเครื่องและส่งข้อมูลดังกล่าวไปยังศูนย์สั่งการจังหวัด เมื่อมีการแจ้งเหตุฉุกเฉินเข้ามาในระบบรับแจ้งเหตุฉุกเฉินของศูนย์สั่งการจังหวัดจะมีการแจ้งเตือนขึ้น เจ้าหน้าที่ประจำศูนย์สั่งการจังหวัดจะใช้ระบบเพื่อตรวจสอบการแจ้งเหตุฉุกเฉิน จากนั้นระบบจะทำการค้นหาว่าตำแหน่งของจุดเกิดเหตุดังกล่าวเกิดขึ้นบนพื้นที่รับผิดชอบของโรงพยาบาลใดจึงจะทำการส่งข้อมูลเหตุฉุกเฉินและตำแหน่งจุดเกิดเหตุไปยังระบบรับแจ้งเหตุฉุกเฉินที่ทำงานบนศูนย์รับแจ้งเหตุฉุกเฉินของโรงพยาบาลนั้น เมื่อศูนย์รับแจ้งเหตุฉุกเฉินของโรงพยาบาลในพื้นที่รับผิดชอบได้รับการแจ้งเหตุฉุกเฉิน ระบบจะทำการแจ้งเตือน จากนั้นเจ้าหน้าที่จะใช้ระบบในการค้นหารถฉุกเฉินของโรงพยาบาลว่าคันใดพร้อมและสามารถเข้าถึงจุดเกิดเหตุได้เร็วที่สุดเพื่อส่งงานไปยังรถคันดังกล่าว โดยจะแสดงลิสต์ (List) รายการรถฉุกเฉินของโรงพยาบาลเรียงตามความพร้อมและความสะดวกในการเข้าช่วยเหลือ เพื่อให้ผู้ใช้งานระบบเลือกรถฉุกเฉินของโรงพยาบาลได้ด้วยความยืดหยุ่น ซึ่งระบบสามารถที่จะแสดงข้อมูลพร้อมทั้งตำแหน่งปัจจุบันของรถฉุกเฉินของโรงพยาบาลทุกคัน รวมทั้งติดตามข้อมูลและสถานะของเหตุฉุกเฉินที่กำลังปฏิบัติงานอยู่ได้



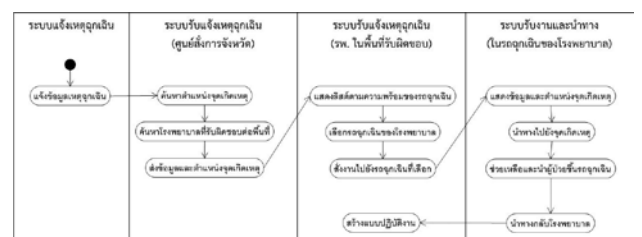
ภาพที่ 1: สถาปัตยกรรมของระบบ

ทั้งนี้ระบบยังสามารถรองรับการแจ้งเหตุฉุกเฉินจากการโทรแจ้งทางโทรศัพท์ ไม่ว่าจะเป็นการโทรแจ้งมาที่ศูนย์สั่งการจังหวัดหรือแจ้งมายังศูนย์รับแจ้งเหตุฉุกเฉินของโรงพยาบาล

โดยตรง ซึ่งจะมีระบบการรับแจ้งเหตุฉุกเฉินทางโทรศัพท์ ด้วยการป้อนข้อมูลในฟอร์มที่ออกแบบให้มีความกระชับและสามารถป้อนข้อมูลได้ด้วยความรวดเร็ว รวมถึงระบบการค้นหาและระบุพิกัดจุดเกิดเหตุบนแผนที่ที่สามารถค้นหาข้อมูลสถานที่จากแหล่งข้อมูลหรือสามารถเลื่อนมาร์คเกอร์ (Marker) เพื่อระบุตำแหน่งที่ใกล้เคียงกับจุดเกิดเหตุที่สุด

เมื่อมีการสั่งงานจากศูนย์รับแจ้งเหตุฉุกเฉินมายังรถฉุกเฉินของโรงพยาบาลที่ถูกเลือก ระบบที่ทำงานอยู่บนรถฉุกเฉินของโรงพยาบาลซึ่งเชื่อมต่อผ่านทางอินเทอร์เน็ตไร้สายและโมดูลของระบบกำหนดตำแหน่งบนโลก จะมีเสียงแจ้งเตือนและแสดงข้อมูลเหตุฉุกเฉินพร้อมทั้งตำแหน่งของจุดเกิดเหตุบนแผนที่ ก่อนที่จะใช้ระบบนำทางไปยังจุดเกิดเหตุเพื่อช่วยในการตัดสินใจของเจ้าหน้าที่ในการเลือกเส้นทางเข้าช่วยเหลือเหตุฉุกเฉิน หากไปถึงยังจุดเกิดเหตุแล้วระบบจะมีการเปลี่ยนการนำทางเป็นเส้นทางไปยังโรงพยาบาล ซึ่งตลอดทั้งการทำงานระบบจะทำการดึงพิกัดตำแหน่งปัจจุบันส่งกลับไปเก็บไว้บนแหล่งเก็บข้อมูลรถฉุกเฉินของโรงพยาบาลเพื่อให้ศูนย์รับแจ้งเหตุฉุกเฉินสามารถติดตามตำแหน่งปัจจุบันของรถฉุกเฉินของโรงพยาบาลได้แบบเรียลไทม์ และมีการบันทึกสถานะของการปฏิบัติงานและบันทึกเวลาในช่วงต่างๆ ของการปฏิบัติหน้าที่เพื่อให้ศูนย์รับแจ้งเหตุฉุกเฉินสามารถติดตามสถานะรวมทั้งนำไปกรอกในแบบบันทึกการปฏิบัติงาน

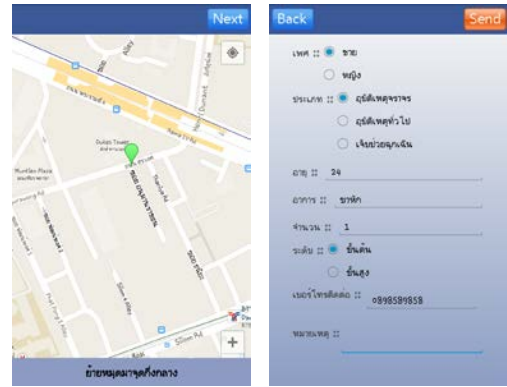
เมื่อการช่วยเหลือเหตุฉุกเฉินเสร็จสิ้น เจ้าหน้าที่จะใช้ระบบสร้างแบบบันทึกการปฏิบัติงาน ซึ่งจะช่วยสร้างแบบบันทึกการปฏิบัติงานได้ทั้งชุดปฏิบัติการฉุกเฉินระดับต้น (BLS) และชุดปฏิบัติการฉุกเฉินระดับกลาง/สูง (ILS/ALS) โดยจะดึงข้อมูลบางส่วนมาจากการทำงานของระบบก่อนหน้า เช่น ข้อมูลผู้ป่วย ข้อมูลผู้ปฏิบัติหน้าที่ ข้อมูลสถานที่ ข้อมูลเวลาในช่วงต่างๆ ของการปฏิบัติหน้าที่ เป็นต้น และบันทึกลงยังฐานข้อมูลแบบบันทึกการปฏิบัติงาน



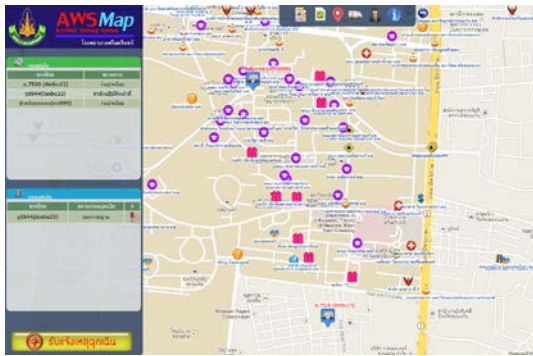
ภาพที่ 2: กระบวนการทำงานของระบบ

2.3.3 การออกแบบส่วนต่อประสานกับผู้ใช้ใช้งาน (User Interface Design)

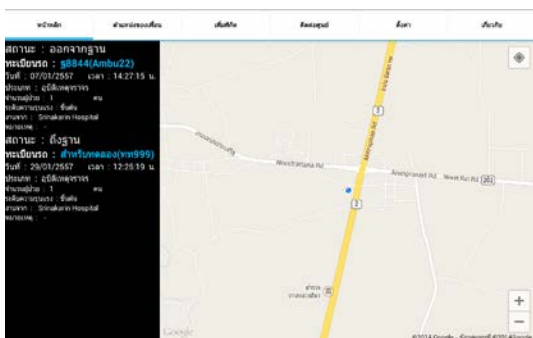
ในการออกแบบส่วนต่อประสานกับผู้ใช้ใช้งานจะเน้นการออกแบบให้ใช้งานง่ายและทำงานอย่างรวดเร็ว โดยระบบการรับแจ้งเหตุฉุกเฉินจะแบ่งส่วนของแผนที่และส่วนของข้อมูลเป็นอย่างสัดส่วน การรับแจ้งเหตุได้จัดการทำงานเป็นขั้นตอน (Step) เพื่อสะดวกต่อการป้อนข้อมูล และมีพื้นที่สำหรับการแสดงแผนที่ขนาดใหญ่ ในส่วนของระบบที่ทำงานบนอุปกรณ์เคลื่อนที่จะใช้แท็บเมนูเพื่อเลือกดูข้อมูลต่างๆ เมื่อมีการรับแจ้งเหตุจะแสดงแผนที่นำทางในขนาดที่สามารถมองเห็นได้อย่างชัดเจน มีปุ่มขนาดใหญ่เพื่อช่วยต่อการสัมผัส



ภาพที่ 5: ส่วนต่อประสานกับผู้ใช้ใช้งานในระบบแจ้งเหตุฉุกเฉินโดยผู้แจ้งเหตุ



ภาพที่ 3: ส่วนต่อประสานกับผู้ใช้ใช้งานในระบบการรับแจ้งเหตุฉุกเฉินที่ศูนย์รับแจ้งเหตุฉุกเฉิน



ภาพที่ 4: ส่วนต่อประสานกับผู้ใช้ใช้งานในระบบรับงานและการนำทาง บนอุปกรณ์เคลื่อนที่แท็บเล็ตบนรถฉุกเฉิน

2.3.4 การพัฒนาระบบ (System Development)

ในขั้นตอนของการพัฒนาระบบจะเป็นการพัฒนาด้วยโปรแกรมประยุกต์ เพื่อทดสอบการทำงานของระบบตามสถาปัตยกรรมที่ได้ออกแบบไว้ว่าสามารถใช้งานได้จริงและอำนวยความสะดวกในการทำงานของเจ้าหน้าที่ปฏิบัติงานช่วยเหลือเหตุฉุกเฉินหรือไม่ ขั้นตอนนี้จะพัฒนาระบบแจ้งเตือนอุบัติเหตุและการระบุตำแหน่งบนแผนที่สำหรับบริการทางการแพทย์ฉุกเฉิน โดยพัฒนาเป็นกรณีศึกษาที่ศึกษาเกี่ยวกับอุบัติเหตุ-ฉุกเฉิน โรงพยาบาลศรีนครินทร์ มหาวิทยาลัยขอนแก่น จังหวัดขอนแก่น ซึ่งงานวิจัยนี้ได้พัฒนาเป็นระบบบนเว็บแอปพลิเคชันและโมบายแอปพลิเคชัน โดยเว็บแอปพลิเคชันพัฒนาด้วยภาษาพีเอชพี (PHP), จาวาสคริปต์ (JavaScript) และใช้เทคนิคเอแจ็กซ์ (Ajax) ในการทำงาน ส่วนโมบายแอปพลิเคชันพัฒนาด้วยภาษาจาวา (Java) ทำงานบนแท็บเล็ตระบบปฏิบัติการแอนดรอยด์ (Android) อ่านพิกัดโดยโมดูลของระบบกำหนดตำแหน่งบนโลก ทั้งสองระบบมีการเชื่อมต่อกันผ่านทางอินเทอร์เน็ต โดยใช้เอ็กซ์เอ็มแอล เป็นตัวกลางของข้อมูลที่จะบันทึกข้อมูลชั่วคราวเท่านั้น ซึ่งเป็นข้อมูลของเหตุฉุกเฉินที่กำลังปฏิบัติงานอยู่ ส่วนฐานข้อมูลแบบบันทึกการปฏิบัติงานจะใช้ฐานข้อมูลมายเอสคิวแอล (MySQL) ในการเก็บข้อมูลแบบบันทึกการปฏิบัติงานทั้งหมดปฏิบัติการฉุกเฉินระดับต้นและชุดปฏิบัติการฉุกเฉินระดับกลาง/สูง จะเก็บข้อมูลเหตุฉุกเฉินทั้งหมดที่ถูกล่วงในการปฏิบัติหน้าที่แล้ว



ภาพที่ 6: บรรยายภาพการทดลองใช้งานระบบ

3. บทสรุปและงานในอนาคต

งานวิจัยนี้ได้พัฒนาระบบแจ้งเตือนอุบัติเหตุและการระบุตำแหน่งบนแผนที่สำหรับบริการทางการแพทย์ฉุกเฉิน ซึ่งมีการสร้างระบบรับแจ้งเหตุฉุกเฉินที่ทำงานบนเว็บแอปพลิเคชันและระบบที่ทำงานบนโมบายแอปพลิเคชัน โดยระบบมีความสามารถครอบคลุมตั้งแต่การแจ้งเหตุฉุกเฉิน การรับแจ้งเหตุฉุกเฉิน จนกระทั่งนำผู้ป่วยเข้ารับการรักษาคาดจนสร้างแบบบันทึกการปฏิบัติงาน จากการนำระบบไปทดลองใช้งานโดยการสร้างสถานการณ์จำลองขึ้น 10 เหตุการณ์ พบว่ารถฉุกเฉินสามารถเข้าถึงจุดเกิดเหตุและนำผู้ป่วยประสบอุบัติเหตุเข้ารับการรักษที่โรงพยาบาลได้เร็วกว่าระบบเดิมเฉลี่ยร้อยละ 41.33 โดยระบบเดิมมีค่าเฉลี่ยของเวลาที่ใช้ในการปฏิบัติงานที่ 20.13 นาที ขณะที่ระบบใหม่ใช้เวลาในการปฏิบัติงานเฉลี่ย 11.81 นาที ซึ่งประเมินจากระยะเวลาตั้งแต่การแจ้งเหตุฉุกเฉิน การรับแจ้งเหตุฉุกเฉิน การวิเคราะห์โรงพยาบาลที่รับผิดชอบในพื้นที่ไปจนถึงจุดเกิดเหตุและออกจากจุดเกิดเหตุ กระทั่งนำผู้ป่วยเข้ารับการรักษายังโรงพยาบาล และจากแบบสอบถามต่อเจ้าหน้าที่ผู้ใช้งานระบบ ด้วยเครื่องมือ โปรแกรมประยุกต์ SPSS ผลคือมีความพึงพอใจในระดับมากต่อการทำงานของระบบ ร้อยละ 75.00 โดยแบ่งเป็น ด้านการรับแจ้งเหตุฉุกเฉินและวิเคราะห์จุดเกิดเหตุมีความพึงพอใจในระดับมาก ร้อยละ 62.50 ด้านระบบการนำทางมีความพึงพอใจในระดับมาก ร้อยละ 75.00 ส่วนในด้านระบบการสร้างและจัดเก็บแบบบันทึกการปฏิบัติงานมีความพึงพอใจในระดับมากที่สุด ร้อยละ 62.50

สำหรับแนวทางการพัฒนาระบบในอนาคตต้องมีการพัฒนาระบบการจัดการรถฉุกเฉินให้รองรับกับเหตุการณ์ที่ซับซ้อนยิ่งขึ้น มีการเชื่อมต่อกับข้อมูลผู้ป่วยของโรงพยาบาล มีการติดตั้งกล้องบนรถฉุกเฉินเพื่อส่งรูปภาพปัจจุบันของผู้ป่วย รวมถึงส่งข้อมูลสัญญาณชีพ เพื่อให้แพทย์เตรียมพร้อมในการรักษาผู้ป่วยได้

4. กิตติกรรมประกาศ

ขอขอบพระคุณ สำนักงานกองทุนสนับสนุนการวิจัย ทุนวิจัยมหบัณฑิต (สกว.) สาขาวิทยาศาสตร์และเทคโนโลยี ภายใต้โครงการเชื่อมโยงภาคการผลิตกับงานวิจัย ทุน สกว.-อุตสาหกรรม (สกว.-สว.) ประจำปี 2555 ที่ได้สนับสนุนงบประมาณในการวิจัย และขอขอบพระคุณ โรงพยาบาลศิริราช นครินทร์ จังหวัดขอนแก่น ซึ่งเป็นสถานประกอบการที่ใช้ในการวิจัยครั้งนี้ ซึ่งความเห็นในรายงานผลการวิจัยเป็นของผู้รับทุน สำนักงานกองทุนสนับสนุนการวิจัย และ โรงพยาบาลศิริราช นครินทร์ ไม่จำเป็นต้องเห็นด้วยเสมอไป

เอกสารอ้างอิง

- [1] World Wide Web Consortium. (2002). "Web Services". [Online], Retrieved April, 2007, from <http://www.w3.org/2002/ws>.
- [2] W3C, "World Wide Web Consortium, Extensible Markup Language (XML)". [Online], Retrieved Jun, 2013, from <http://www.w3.org/XML>.
- [3] Google Developer, "Google Map API". [Online], Retrieved Jun, 2013, from <https://developers.google.com/maps/>.
- [4] JM. Zogg, "GPS Basics Introduction to the system Application overview" U-Blox ag, Mar 2002.
- [5] D. Kim, J. Kim, DK. Kim, M. Choa, SK. Yoo, "Real-time Ambulance Location Monitoring using GPS and Maps Open API" *30th Annual International IEEE EMBS Conference*, Vancouver, BC, Aug 20-25, 2008, pp. 1561-1563.
- [6] SE. Masri, B. Saddik, "Mobile Emergency System and integration" *12th International Conference on Mobile Data Management (MDM)*, Lulea ,June 6-9, 2011, Vol.2, pp.67-72.

แนวทางการทดสอบเรสท์ฟูลเว็บเซอร์วิสในการพัฒนาแบบเอไจล์

An Approach to Testing RESTful Web Services in Agile Development

สรिता คุปตยานนท์ (Sarita Khuptayanon) และ ญาใจ ลิ้มปิยะกรณ์ (Yachai Limpiyakorn)

ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

Sarita.Kh@student.chula.ac.th, Yachai.L@chula.ac.th

บทคัดย่อ

เซอร์วิสต่างๆ ที่ให้บริการส่วนมากบนแอปพลิเคชันมือถือที่กำลังเติบโตอยู่ในขณะนี้ มักนิยมพัฒนาเป็นเรสท์ฟูลเว็บเซอร์วิส เนื่องจากขนาดข้อมูลที่รับส่งผ่านเรสท์ฟูลเว็บเซอร์วิสมีขนาดเล็ก ระเบียบวิธีพัฒนาซอฟต์แวร์แบบเอไจล์เหมาะสมกับการพัฒนาแอปพลิเคชันมือถือที่มีการเปลี่ยนแปลงความต้องการรวดเร็ว ด้วยเทคนิคการพัฒนาขับเคลื่อนด้วยการทดสอบ กรณีทดสอบจะถูกพัฒนาก่อนการเขียนรหัสชุดคำสั่งในทุกๆ รอบของการพัฒนา ถ้าหากมีเครื่องมือสนับสนุนให้การทดสอบสามารถกระทำได้อย่างอัตโนมัติ จะส่งผลให้สามารถพัฒนาบริการต่างๆ ได้มากและเร็วยิ่งขึ้น รวมถึงสามารถลดค่าใช้จ่ายในการทดสอบ งานวิจัยนี้จึงนำเสนอแนวทางและพัฒนาระบบเพื่อสนับสนุนการทดสอบเรสท์ฟูลเว็บเซอร์วิส นอกจากนี้ กรณีทดสอบที่สร้างขึ้นแบบอัตโนมัติในบริบทการพัฒนาซอฟต์แวร์แบบเอไจล์ในงานวิจัยนี้ สามารถนำกลับมาใช้ใหม่สำหรับการทดสอบในลักษณะอื่นๆ ได้ในภายหลัง

คำสำคัญ: การสร้างกรณีทดสอบ เรสท์ฟูลเว็บเซอร์วิส เอไจล์ ความสามารถนำกลับมาใช้ใหม่

Abstract

Today various growing services provided as mobile applications are often developed with RESTful web services as the size of data transferred over the RESTful web services is small. Agile software development methodology suits for the development of mobile applications with rapid changing requirements. When

using the technique of Test -Driven Development, the test cases will be developed prior to coding during each cycle of development or sprint. If a tool exists to support automated testing, it will lead to more and faster development of these services. The cost of testing will be reduced as well. This research thus presents an approach and develops a system to support the RESTful web service test. Moreover, the test cases automatically generated in the context of Agile software development presented in this research work can be reused later for other types of testing.

Keyword: test case generation, RESTful web services, agile, reusability.

1. บทนำ

การสื่อสารทางอินเทอร์เน็ตทำให้เกิดระบบกระจาย (Distributed system) ซึ่งได้ให้บริการและข้อมูลต่างๆ อยู่มากมาย ทั้งในรูปแบบที่เป็น Public และ Private ระบบกระจายต่างๆ เหล่านี้ สามารถเรียกใช้เซอร์วิสและเข้าถึงข้อมูลซึ่งกันและกัน ทำให้เป็นเรื่องยากที่จะสามารถทดสอบว่าระบบหนึ่งๆ ทำงานได้ผลลัพธ์ตามที่ต้องการได้โดยสมบูรณ์ อันเนื่องมาจากปัจจัยหลายประการ ที่เห็นเด่นชัดที่สุด คือ ความยากในการทดสอบเซอร์วิสต่างๆ ที่ระบบเหล่านี้ให้บริการ

การพัฒนาซอฟต์แวร์แบบเอไจล์ (Agile) ด้วยเทคนิคการพัฒนาขับเคลื่อนด้วยการทดสอบ (Test-Driven Development) จะต้องมีการทดสอบทุกๆ รอบของการพัฒนา

ซอฟต์แวร์ที่เรียกว่า Sprint ซึ่งในแต่ละรอบของการพัฒนาซอฟต์แวร์ นักพัฒนาจะสร้างกรณีทดสอบจาก User story ก่อนการเขียนโค้ดและทำการทดสอบเอง แต่สุดท้ายแล้ว ต้องให้นักทดสอบ (Tester) มาทดสอบอีกครั้ง เนื่องจากในมุมมองของนักพัฒนา อาจจะไม่สามารถทดสอบได้ครบถ้วนทุกกรณี ดังนั้น เมื่อสิ้นสุดในแต่ละรอบของการพัฒนาซอฟต์แวร์แล้ว นักทดสอบจึงต้องทดสอบทุกๆ ฟังก์ชันหรือเซอร์วิสใหม่หมดทุกครั้ง หากมีข้อผิดพลาดหรือมีการเปลี่ยนแปลงความต้องการกะทันหัน กรณีทดสอบชุดเดิมอาจจะต้องถูกนำมาใช้อีกครั้งเพื่อทดสอบในกรณีดังกล่าว หรือเมื่อเซอร์วิสของผู้ให้บริการมีการเปลี่ยน API ก็สามารถทดสอบได้อย่างรวดเร็วว่าส่งผลกระทบต่อการพัฒนาหรือไม่ ด้วยเหตุผลทั้งหลายดังกล่าวข้างต้น หากมีเครื่องมือสนับสนุนให้การทดสอบที่เรียกใช้เรสท์ฟูลเว็บเซอร์วิส (RESTful web services) สามารถกระทำได้อย่างรวดเร็ว จะเป็นการขับเคลื่อนธุรกิจองค์กรให้สามารถพัฒนาบริการต่างๆ ได้มากและเร็วยิ่งขึ้น รวมทั้งสามารถตอบสนองต่อธุรกิจที่ผลิตภัณฑ์มีการเปลี่ยนแปลงอย่างรวดเร็ว เช่น แอปพลิเคชันบนมือถือ เป็นต้น

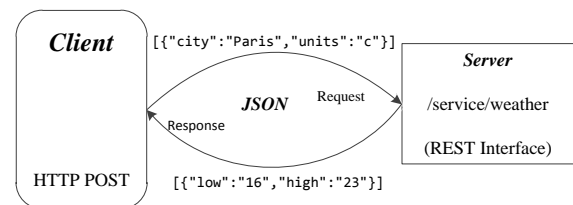
2. วรรณกรรมที่เกี่ยวข้อง

2.1 เรสท์ (REST : Representational State Transfer)

เรสท์ [1] เป็นสถาปัตยกรรมแบบหนึ่งของระบบเน็ตเวิร์ก โดยเรียกใช้ผ่านทางโปรโตคอลเอชทีทีพี (HTTP) ในการแลกเปลี่ยนข้อมูลผ่านเมธอด (method) ได้แก่ GET/ POST/ PUT/ DELETE และส่งออกข้อมูลในรูปแบบต่างๆ ได้ ทำให้ปริมาณข้อมูลที่รับส่งน้อยกว่าการใช้โปรโตคอล SOAP อยู่มาก ซึ่งเป็นข้อดีของเรสท์ที่ทำให้เป็นที่นิยม เพราะมีผลกับประสิทธิภาพของการใช้งานโปรแกรมอีกด้วย นอกจากนี้ผู้พัฒนาเว็บเซอร์วิสสามารถสร้างและเรียกใช้เว็บเซอร์วิสได้ง่าย เพราะเลือกใช้เพียงเอ็กซ์เอ็มแอล (XML: eXtensible Markup Language) และเอชทีทีพี

2.2 เจสัน (JSON : JavaScript Object Notation)

เจสัน [2] หรือ JavaScript Object เป็นรูปแบบการแลกเปลี่ยนข้อมูลระหว่างไคลเอ็นต์และเซิร์ฟเวอร์ที่ได้รับความนิยมมาก พัฒนาโดย Douglas Crockford ข้อมูลในรูปแบบของเจสันเป็นข้อความธรรมดา (Plain text) ที่ถูกห่อหุ้มด้วยเครื่องหมายปีกกา { } และถูกจัดเรียงให้อยู่ในรูปแบบ Object Array เจสันสนับสนุนข้อมูล 6 ประเภท ได้แก่ Strings, Number, Booleans, Arrays, Object และ Null ปัจจุบันเจสันได้รับความนิยม เนื่องจากกระชับ เข้าใจง่าย และสามารถใช้งานได้ในทุกภาษาที่เป็นนิยมน อีกทั้งยังมีไลบรารีของภาษาโปรแกรมอื่นๆ ที่ใช้ประมวลผลข้อมูลในรูปแบบเจสันมากมาย ด้วยเหตุนี้งานวิจัยจึงสนับสนุนการทดสอบเรสท์ฟูลเว็บเซอร์วิส ที่มีรูปแบบการแลกเปลี่ยนข้อมูลแบบเจสัน ภาพที่ 1 แสดงตัวอย่างรูปแบบเจสันที่ใช้ในการรับส่งข้อมูลบนเครือข่าย



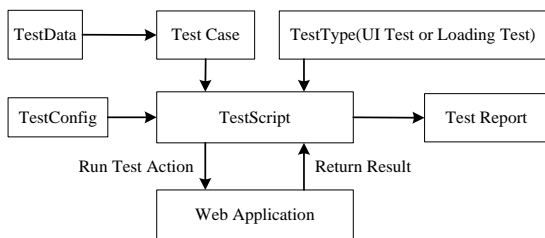
ภาพที่ 1: ตัวอย่างข้อมูลในรูปแบบเจสันที่ใช้ในการติดต่อบนเครือข่าย

2.3 A Test Automation Framework Based on WEB [3]

บทความนี้ได้กล่าวถึงการออกแบบกรอบงานทดสอบแบบอัตโนมัติสำหรับเว็บแอปพลิเคชัน เพื่อให้มีความสะดวกในการทดสอบฟังก์ชัน, การทดสอบประสิทธิภาพ, การนำกลับมาใช้ใหม่ และตรวจสอบความถูกต้อง โดยใช้เครื่องมือทดสอบ Selenium เพื่อแก้ปัญหาการใช้งานเว็บด้วยเบราว์เซอร์ที่หลากหลาย และใช้เครื่องมือ Jmeter ในการสนับสนุนการร้องขอไปยังเว็บเซิร์ฟเวอร์สำหรับทดสอบ โดยงานวิจัยได้ออกแบบกรอบงานทดสอบแบบอัตโนมัติ แสดงดังภาพที่ 2 โดยสามารถนำมาประยุกต์ใช้ในงานวิจัยได้

2.4 Test-the-REST: An Approach to Testing RESTful Web Service [4]

งานวิจัยดังกล่าวได้นำเสนอวิธีการทดสอบเรสท์ฟูลเว็บเซอร์วิส โดยใช้รูปแบบของเอ็กซ์เอ็มแอลมาช่วยในการสร้างและตรวจสอบความถูกต้องของกรณีทดสอบ รวมถึงการนำผลลัพธ์ของการทดสอบมาแสดงผลด้วย โดยในงานวิจัยได้นำวิธีการดังกล่าวมาประยุกต์ใช้ด้วย ข้อดีของการใช้เอ็กซ์เอ็มแอลคือ ง่ายต่อการเรียกข้อมูล หรือเพิ่มข้อมูล ภาพที่ 3 แสดงตัวอย่างกรณีทดสอบ 2 กรณีในรูปแบบเอ็กซ์เอ็มแอลที่จะถูกเรียกเพื่อประมวลผลตามลำดับที่กำหนดในสคริปต์ทดสอบ (Test Script) ดังแสดงในภาพที่ 4



ภาพที่ 2: A web-based test automation framework [3]

```

<testcase>
<id>TTR-BL-1</id>
<URI>http://booklisthost.com/books</URI>
<atomicity>atomic</atomicity>
<method>POST</method>
<input>War and Peace|Leo Tolstoy</input>
</testcase>

<testcase>
<id>TTR-BL-2</id>
<URI>http://booklisthost.com/book/book1
</URI>
<atomicity>atomic</atomicity>
<method>GET</method>
</testcase>
    
```

ภาพที่ 3: ตัวอย่างกรณีทดสอบในรูปแบบเอ็กซ์เอ็มแอล [4]

```

<testcase>
<id>TTR-BL-3</id>
<atomicity>composite</atomicity>
<sequence>
<testcase>
<src>TTR-BL-1.xml</src>
</testcase>
<testcase>
<src>TTR-BL-2.xml</src>
</testcase>
</sequence>
</testcase>
    
```

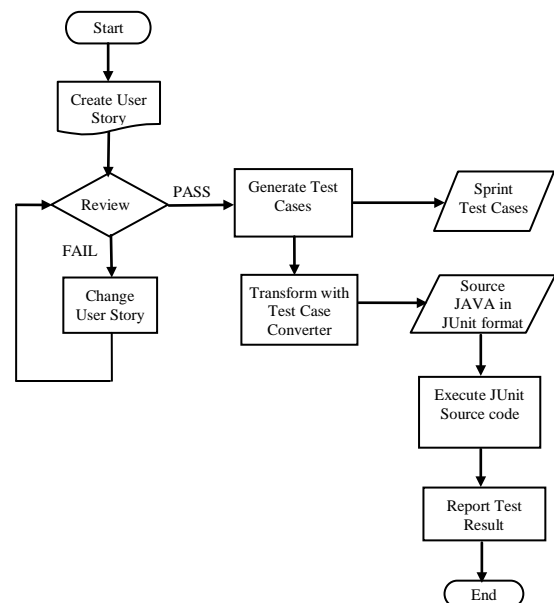
ภาพที่ 4: ตัวอย่างสคริปต์ทดสอบในรูปแบบเอ็กซ์เอ็มแอล [4]

2.5 JUnit

JUnit [5] เป็นกรอบงานทดสอบ (Testing framework) พัฒนาโดย Kent Beck และ Erich Gamma มีจุดประสงค์เพื่อให้เป็นกรอบงานการทดสอบหน่วย (Unit Testing) สำหรับโปรแกรมภาษาจาวา โดย JUnit จะมี Assert แบบต่างๆ กำหนดอยู่ใน class org.junit.Assert ข้อดีของ Assert คือ มีคุณสมบัติมากมายที่ทำให้ง่ายต่อการเขียนและเรียกใช้การทดสอบ แต่ละ Assert มีความเหมาะสมที่จะใช้ในงานวิจัยที่ใช้ทดสอบการทำงานของเรสท์ฟูลเว็บเซอร์วิส นอกจากนี้ ข้อสัปดาห์ของ JUnit มีความยืดหยุ่น สามารถปรับแก้ หรือสามารถนำกลับมาใช้ซ้ำได้

3. วิธีการดำเนินการวิจัย

งานวิจัยนี้แนะนำเสนอแนวทางในการทดสอบซอฟต์แวร์ซึ่งมีการเรียกใช้เซอร์วิสที่พัฒนาด้วยเรสท์ ประกอบด้วยขั้นตอนต่างๆ ดังภาพที่ 5 เริ่มต้นจากความต้องการของซอฟต์แวร์จะถูกแบ่งออกเป็นส่วนย่อยๆ เรียกว่า User story โดยความต้องการที่สำคัญจะถูกเลือกมาพัฒนาในอันดับแรกๆ



ภาพที่ 5: ขั้นตอนวิธีการทำงานของระบบ

3.1 Create User Story

การเขียนความต้องการของระบบ หรือ User story ควรอยู่ในรูปแบบประโยคที่เข้าใจได้ง่าย สามารถเป็นตัวกลางสื่อสารกันระหว่างทีมพัฒนาซอฟต์แวร์กับลูกค้าหรือผู้ใช้งาน และนำไปสู่การยืนยันความต้องการที่ตรงกันของทั้งสองฝ่าย ในงานวิจัยนี้ได้กำหนดให้เขียน User story ในรูปแบบ Excel ตามช่องที่กำหนดไว้ และต้องใส่ค่าทุกช่อง เพื่อให้ง่ายต่อการตรวจสอบและสามารถสร้างชุดทดสอบได้อัตโนมัติ

ตารางที่ 1 แสดงตัวอย่างรูปแบบ Excel ที่กำหนดขึ้นเพื่อบรรยาย User story ที่มีการเรียกใช้เรสทูล์เว็บเซอร์วิส ในขั้นตอนนี้นำเป็นตัวอย่างการประกอบกันของนักทดสอบมาร่วมสร้าง User story ด้วย เนื่องจากจะต้องมีการกำหนดเกณฑ์การยอมรับ ซึ่งประกอบด้วยกรณีทดสอบที่หลากหลายเงื่อนไข และอยู่ในรูปแบบตามที่กำหนด เพื่อง่ายต่อการนำไปใช้ในขั้นตอนถัดไป

ตารางที่ 1: รูปแบบการเขียน User story ที่นำเสนอในงานวิจัย

Feature	Withdraw			
Description	Customers enter the account number and the withdrawal amount.			
API Name	withdraw			
Request URI	http://localhost:8282/bank/withdraw			
Request Method	GET			
Request attribute	account number	Data type	String	
	amount	Data type	Decimal	
Response attribute	result code	Data type	Integer	
TEST-CASE- ID-001				Assertion method
success	input	account number	2692874928	Equals
		amount	2000	Equals
	output	result	0	Equals
TEST-CASE- ID-002				Assertion method
insufficient	input	account number	2692874928	Equals
		amount	30000	Equals
	output	result	1	Equals
TEST-CASE- ID-003				Assertion method
invalid account number	input	account number	8920192211	Equals
		amount	2000	Equals
	output	result	2	Equals

คุณลักษณะต่างๆใน User story ประกอบด้วย

- Feature: ชื่อคุณสมบัติของซอฟต์แวร์
- Description: คำอธิบายโดยสังเขปของ User Story

- API Name: ชื่อช่องทางการเชื่อมต่อซึ่งเป็นตัวกลางที่ทำให้โปรแกรมประยุกต์เชื่อมต่อกับโปรแกรมประยุกต์อื่น
- Request URI: Uniform resource identifier ที่ร้องขอ
- Request Method: ได้แก่ GET/ POST/ PUT/ DELETE
- Request attribute: ค่าที่จำเป็นต้องส่งไป
- Response attribute: ค่าผลลัพธ์
- Data type: ประเภทของข้อมูล
- TEST-CASE- ID-XXX: Test case ID
- Assertion method: การเปรียบเทียบกับค่าที่จะเกิดขึ้นจริง ประกอบด้วย 8 ค่า คือ ArrayEquals, Equals, Null, NotNull, Same, NotSame, Equals, That

3.2 Review

การตรวจทานความถูกต้องสมบูรณ์ของ User Story ประกอบด้วย 2 ขั้นตอน คือ 1) ตรวจทานโดยอาศัยความรู้ความชำนาญเฉพาะด้านของนักทดสอบ และ 2) ตรวจทานความถูกต้องสมบูรณ์ของข้อมูล เช่น ค่าทุกช่องห้ามว่าง และประเภทข้อมูลของค่าถูกต้องเหมาะสม เป็นต้น

3.3 Change User Story

กระบวนการเปลี่ยนแปลงหรือแก้ไข User Story ภายหลังจากพบข้อผิดพลาดในขั้นตอนการตรวจทาน จะถูกดำเนินการโดยผู้ที่มีส่วนเกี่ยวข้องกับซอฟต์แวร์ ได้แก่ นักพัฒนา นักทดสอบ นักออกแบบ ผู้เชี่ยวชาญด้านเทคโนโลยีสารสนเทศ และผู้จัดการโครงการ

3.4 Generate Test Cases

เป็นกระบวนการแปลง User story ที่ผ่านการตรวจทานแล้วให้เป็นกรณีทดสอบที่อยู่ในรูปแบบเอ็กซ์เอ็มแอล สามารถสรุปเป็นขั้นตอนได้ดังนี้

- 1) อ่านไฟล์ Excel โดยใช้ Apache POI
- 2) สร้างไฟล์ xsd และสร้าง Class ที่เกี่ยวข้อง
- 3) อ่านค่าในแถวแต่ละแถว และทำการเรียกใช้ค่าผ่านเมธอด getter/setter โดยสร้างวัตถุที่ตรงกันไว้ใน Class ที่เกี่ยวข้อง
- 4) เก็บค่าลง Arraylist

5) ใช้ JAXB Architecture Marshalling ส่งค่า ArrayList ไปยังเพิ่มผลลัพธ์เอ็กซ์เอ็มแอล

ตัวอย่างไฟล์ผลลัพธ์กรณีทดสอบในรูปแบบเอ็กซ์เอ็มแอล ดังแสดงในภาพที่ 6

```
<testcase>
<id>TEST-CASE-ID-01</id>
<URI>http://localhost:8282/bank/withdraw</URI>
<method>GET</method>
<request>
<param account_number="2692874928" assert="Equals"/>
<param amount="2000" assert="Equals"/>
</request>
<response>
<param result="0" assert="Equals"/>
</response>
</testcase>
```

ภาพที่ 6: ตัวอย่างกรณีทดสอบในรูปแบบเอ็กซ์เอ็มแอล

3.5 Transform with Test Case Converter

กระบวนการนี้ทำการแปลงกรณีทดสอบในรูปแบบเอ็กซ์เอ็มแอลจากขั้นตอนก่อนหน้าให้เป็นสคริปต์ทดสอบที่อยู่ในรูปแบบ JUnit source code ตัวอย่างดังภาพที่ 7 ซึ่งสะดวกต่อการนำไปใช้งานผ่าน IDE Tools ต่างๆ โดยเครื่องมือเหล่านี้ส่วนมากจะสนับสนุน JUnit

สคริปต์ทดสอบที่ได้สามารถนำกลับมาใช้ทดสอบได้บ่อยครั้งเท่าที่ต้องการ โดยเวลาทดสอบจะแยกทดสอบเฉพาะส่วน Request attribute และ Response attribute ของ API Name หนึ่งๆ เท่านั้น ทั้งนี้ สคริปต์ทดสอบจะมุ่งเน้นใช้ในการทดสอบความพร้อมของเรสต์ฟูลเว็บเซอร์วิส ก่อนที่เริ่มจะเข้าพัฒนาซอฟต์แวร์ ข้อดีคือ การแบ่งการทดสอบเป็นส่วนย่อยๆ ทำให้ปรับเปลี่ยนแก้ไขได้ง่ายหากซอฟต์แวร์มีความผิดพลาดหรือพัฒนาแล้วไม่เป็นไปตามความต้องการ และด้วยลักษณะการพัฒนาที่เน้นให้ซอฟต์แวร์ผ่านการทดสอบก่อน แล้วนำมาปรับเปลี่ยนภายหลัง ทำให้ลดระยะเวลาการพัฒนา สามารถพัฒนาส่วนต่อไปได้ ในขณะที่ส่วนแรกกำลังทำการปรับปรุงเซอร์วิสหรือซอฟต์แวร์

3.6 Execute JUnit Source Code

ทำการประมวลผล JUnit Source Code เพื่อทดสอบเรสต์ฟูลเว็บเซอร์วิสที่ต้องการ กล่าวคือ ตรวจสอบความถูกต้องในการเรียกใช้แต่ละ API และตรวจสอบความถูกต้อง Assertion methods ตามข้อมูลที่ระบุจาก User Story การประมวลผลจะใช้ Maven ซึ่งเป็นเครื่องมือการจัดสร้างโครงสร้างของซอฟต์แวร์ และสามารถดาวน์โหลด Java Class Library ที่ต้องการสำหรับการประมวลผลมาให้ โดยผู้ใช้งานจะต้องกำหนดค่า JUnit ในไฟล์ pom.xml ก่อน รวมทั้งระบุเวอร์ชันที่ต้องการ ดังภาพที่ 8 JUnit Library นี้จะทำการประมวลผล JUnit ให้อยู่ในรูปแบบเจสัน ซึ่งมีขนาดข้อมูลเล็กและเหมาะสมสำหรับการทดสอบเรสต์ฟูลเว็บเซอร์วิส

```
@Before
public void setUp() {
controller=new DirectWithdrawController();
withdrawService=mockito.mock(WithdrawService.class);
controller.setDirectWithdrawService(withdrawService);
confirmRequest=new ConfirmRequest();
confirmResponse=new ConfirmResponse();
confirmRequest.setAccountNumber("2692874928");
confirmRequest.setAmount(new BigDecimal(2000));
confirmResponse.setResult("0");
}
```

ภาพที่ 7: ตัวอย่าง JUnit source code

```
<dependencies>
[... ]
<dependency>
<groupId>junit</groupId>
<artifactId>junit</artifactId>
<version>4.8.1</version>
<scope>test</scope>
</dependency>
[... ]
</dependencies>
```

ภาพที่ 8: ค่าที่กำหนดใน pom.xml

3.7 Report Test Result

ทำการเปรียบเทียบผลลัพธ์ที่ได้จากการประมวลผล JUnit Source Code กับค่า Response attribute ที่ระบุใน User story และสร้างค่าลงใน XML Tag Element <test_result>{result test}</test_result> มีค่าเป็น PASS หรือ FAIL ดังแสดงในภาพที่ 9 เพื่อรอการอ่านค่าและกลับแสดงผลที่หน้าจอ โดยการอ่านค่าที่เอ็กซ์เอ็มแอล จะประมวลผลในขั้นตอนนี้ด้วย

4. ผลการดำเนินการวิจัย

ในส่วนนี้จะได้นำเสนอตัวอย่างหน้าจอการทดสอบเรสต์ฟูลเว็บเซอร์วิสด้วยขั้นตอนวิธีที่นำเสนอ ภาพที่ 10 แสดงตัวอย่างหน้าจอโปรแกรมผลลัพธ์การทดสอบเรสต์ฟูลเว็บเซอร์วิส คอลัมน์ ID แสดงลำดับสคริปต์ที่ทดสอบ และคอลัมน์ Test Result แสดงผลลัพธ์การทดสอบเป็น PASS หรือ FAIL จะแสดง FAIL ก็ต่อเมื่อเรสต์ฟูลเว็บเซอร์วิสทำงานผิดพลาด ทำให้ผลลัพธ์ที่ตอบกลับไม่ตรงกับที่ระบุใน User Story

```
<testcase>
<id>TEST-CASE- ID-01</id>
<URI>http://localhost:8282/bank/withdraw</URI>
<method>GET</method>
<request>
<param account_number="2692874928"
assert="Equals" />
<param amount="2000" assert="Equals" />
</request>
<response>
<param result="0" assert="Equals" />
</response>
<test_result>PASS</test_result>
</testcase>
```

ภาพที่ 9: ตัวอย่างผลลัพธ์ไฟล์เอ็กซ์เอ็มแอลจากการประมวลผล JUnit source code

Test RESTful Web Services

User Story...

Browse
Send

ID	Test Result
TEST-CASE-ID-001	PASS
TEST-CASE-ID-002	PASS
TEST-CASE-ID-003	FAIL
TEST-CASE-ID-004	PASS
TEST-CASE-ID-005	FAIL

ภาพที่ 10: ตัวอย่างหน้าจอ โปรแกรมการทดสอบเรสต์ฟูลเว็บเซอร์วิส

5. สรุป

งานวิจัยนี้ได้พัฒนาระบบเพื่ออำนวยความสะดวกการทดสอบเรสต์ฟูลเว็บเซอร์วิส ในบริบทการพัฒนาซอฟต์แวร์แบบเอไจล์ที่ใช้เทคนิคการพัฒนาขับเคลื่อนด้วยการทดสอบ ซึ่งเหมาะสมกับการพัฒนาซอฟต์แวร์ที่มีการเปลี่ยนแปลงความต้องการรวดเร็วและสามารถแบ่งแยกออกเป็นความต้องการย่อยๆได้ ระบบที่พัฒนาขึ้นสามารถสร้างกรณีทดสอบอัตโนมัติในรูปแบบเอ็กซ์เอ็มแอล จาก User story ที่เขียนตามข้อกำหนดที่ระบุไว้ในงานวิจัย กรณีทดสอบเอ็กซ์เอ็มแอลจะถูกแปลงเป็น JUnit source code ซึ่งสามารถนำกลับมาใช้ทดสอบซ้ำได้ หากมีการเปลี่ยนแปลงความต้องการหรือ โปรแกรมมีความผิดพลาด นอกจากนี้ รูปแบบข้อมูลเจสันที่ใช้ส่งผ่านข้อมูลบนเครือข่ายนั้น มีขนาดเล็ก ทำให้เพิ่มประสิทธิภาพในการทดสอบ วิธีการที่นำเสนอช่วยให้การพัฒนาเชื่อมต่อเซอร์วิสต่างๆเข้าด้วยกันเป็นไปได้ง่ายขึ้น สามารถตรวจสอบการทำงานเรสต์ฟูลเว็บเซอร์วิสได้บ่อยเท่าที่ต้องการว่าเป็นไปตามผลลัพธ์ที่คาดหวังหรือไม่ ก่อนที่จะเริ่มพัฒนาซอฟต์แวร์ในแต่ละรอบการพัฒนา รวมทั้งช่วยลดความซ้ำซ้อนในการทดสอบ ทำให้ลดทรัพยากรที่ใช้ในการทดสอบได้

เอกสารอ้างอิง

- [1] RESTful Web services: The basic [อินเทอร์เน็ต]. [อ้างถึง 20 กรกฎาคม 2013]. สืบค้น จาก: <http://www.ibm.com/developerworks/webservices/library/ws-restful/>
- [2] S. Srinivas Sriparasa, *JavaScript and JSON Essentials*, Packt, Birmingham: UK, 2013.
- [3] F. Wang and W. Du, "A Test Automation Framework Based on WEB," *Proceedings of the 11th International Conference on Computer and Information Science 2012 (ICIS)*, May 2012–June 2012, pp. 638–687.
- [4] S. Kumar Chakrabarti and P. Kumar, "Test-the-REST: An Approach to Testing RESTful Web-Services," *Computation World 2009*, Nov 2009, pp.302–308.

- [5] P. Tahchiev, F. Leme, V. Massol, and G. Gregory, *JUnit IN ACTION*, Manning, Stamford: CT, 2011.

การพัฒนาเว็บแอปพลิเคชันบนแพลตฟอร์มกูเกิลแอปเอนจินแบบอิงแบบจำลอง

A Model-Driven Development of Web-Based Applications on Google App Engine Platform

สมฤดี แก้วขาว (Somrudee Kaewkao) และ ทวีติย์ เสนีวงศ์ ณ อยุธยา (Twittie Senivongse)

ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

somrudee.ka@student.chula.ac.th, twittie.s@chula.ac.th

บทคัดย่อ

งานวิจัยนี้นำเสนอการพัฒนาเว็บแอปพลิเคชันบนแพลตฟอร์มกูเกิลแอปเอนจินด้วยแนวคิดของเอ็มดีเอ โดยการสร้างแบบจำลองในระดับที่เฉพาะเจาะจงกับแพลตฟอร์ม แล้วทำการแปลงเป็นโค้ดเทมเพลตสำหรับเว็บแอปพลิเคชัน สำหรับให้นักพัฒนาทำการเติมให้เป็นโค้ดที่สมบูรณ์ต่อไป การสร้างแบบจำลองทำโดยอาศัยยูเอ็มแอลโปรไฟล์ WebGAE ที่นำเสนอสำหรับเว็บแอปพลิเคชันบนแพลตฟอร์มกูเกิลแอปเอนจิน จากนั้นใช้กฎการแปลงที่ผู้วิจัยกำหนดเพื่อทำการแปลงแบบจำลองไปเป็นโค้ดเทมเพลตที่ใช้เทคโนโลยีจาวา ผู้วิจัยได้พัฒนาเครื่องมือสำหรับการแปลง รวมถึงนำเสนอกรณีศึกษาที่นำยูเอ็มแอลโปรไฟล์ที่เสนอไปประยุกต์ใช้งาน เพื่อประเมินประสิทธิภาพของการแปลง ผลการประเมินพบว่าวิธีการที่เสนอมีสัดส่วนการแปลงโค้ดได้ประมาณ 45 – 79% ของโค้ดที่สมบูรณ์ จึงสามารถช่วยนักพัฒนาในการพัฒนาเว็บแอปพลิเคชันบนแพลตฟอร์มกูเกิลแอปเอนจินได้สะดวกขึ้น

คำสำคัญ: กูเกิลแอปเอนจิน เอ็มดีเอ ยูเอ็มแอลโปรไฟล์ กฎการแปลง

Abstract

This research proposes a development of Web-based applications for the Google App Engine platform using a Model Driven Architecture (MDA). That is, a platform-specific model will be built for the application and the model will be transformed into the Java application code template so that a developer can later fill in the template to make a complete application. We propose a UML profile for Web applications on Google App Engine, or WebGAE, and a set of transformation rules for the

generation of the code template from the model. We develop a transformation tool and use it to build a case study application to evaluate the efficiency of the transformation. The evaluation result shows that the proposed method has a transformation ratio of 45-79% of the final code completed by the developer. Hence this method can help facilitate the developers in constructing Web-based applications on Google App Engine.

Keyword: Google App Engine, MDA, UML Profile, Transformation Rules

1. บทนำ

ในปัจจุบันการประมวลผลแบบคลาวด์ (Cloud Computing) กลายเป็นทางเลือกหนึ่งสำหรับองค์กรที่ต้องการประหยัดและลดภาระการจัดการด้านระบบสารสนเทศ เนื่องจากองค์กรสามารถเลือกซื้อหรือใช้บริการจากผู้ให้บริการคลาวด์ ทั้งในด้านโครงสร้างพื้นฐาน แพลตฟอร์มการพัฒนาระบบ และแอปพลิเคชัน โดยไม่ต้องลงทุนด้านฮาร์ดแวร์หรือซอฟต์แวร์เอง กูเกิลแอปเอนจิน [1] เป็นแพลตฟอร์มหนึ่งที่ได้รับนิยมนิยมสำหรับการพัฒนาแอปพลิเคชันเพื่อให้ทำงานบนโครงสร้างพื้นฐานของกูเกิล โดยผู้พัฒนาทำการเรียนรู้เกี่ยวกับเอพีไอ ไลบรารี และเครื่องมือต่าง ๆ ของกูเกิลแอปเอนจิน จากนั้นจะสามารถพัฒนาและดีพลอยแอปพลิเคชันได้โดยไม่ต้องมีโครงสร้างพื้นฐานของตัวเอง

ผู้วิจัยเห็นว่าการพัฒนาแอปพลิเคชันบนกูเกิลแอปเอนจินนั้น หากมีเครื่องมือที่ช่วยในการสร้างโค้ดของแอปพลิเคชันก็จะช่วยให้การพัฒนาทำได้สะดวกยิ่งขึ้น แนวคิดนี้เป็นไปตามแนวคิดของสถาปัตยกรรมอิงแบบจำลองหรือเอ็มดีเอ

(Model-Driven Architecture: MDA) [2] ของโอเอ็มจี (OMG) ซึ่งเห็นว่าผู้พัฒนาควรมุ่งเน้นที่การออกแบบแบบจำลองของระบบแอปพลิเคชันให้มีคุณภาพดีและทำงานตามที่ต้องการได้ ส่วนการพัฒนาโค้ดตามการออกแบบนั้นสามารถใช้เครื่องมือแปลงจากแบบจำลอง โดยมีการกำหนดกฎการแปลง (Transformation Rules) ไว้แนวทางนี้จะช่วยลดปริมาณงานที่นักพัฒนาต้องทำในการพัฒนาแอปพลิเคชัน อีกทั้งยังสามารถนำแบบจำลองและกฎการแปลงมาใช้ซ้ำสำหรับการแปลงหลากหลายแอปพลิเคชันบนหลากหลายแพลตฟอร์มการทำงานได้ เอ็มดีเอเสนอการแปลงระหว่างแบบจำลองของซอฟต์แวร์สามระดับได้แก่ 1) แบบจำลองที่อิสระจากแพลตฟอร์ม (Platform Independent Model: PIM) สำหรับใช้อธิบายถึงการทำงานของระบบตามโดเมนของธุรกิจ 2) แบบจำลองที่เฉพาะเจาะจงกับแพลตฟอร์ม (Platform Specific Model: PSM) เป็นแบบจำลองของระบบตามโดเมนของธุรกิจที่อยู่บนแพลตฟอร์มหนึ่ง ๆ ของการทำงาน และ 3) โค้ด (Code) ถือเป็นแบบจำลองในรูปของโปรแกรมตามโดเมนของธุรกิจซึ่งเขียนตามภาษาและเทคโนโลยีของแพลตฟอร์มที่เลือกใช้ กฎการแปลงจะถูกกำหนดขึ้นเพื่อแปลงพีไอเอ็มเป็นพีเอสเอ็มและแปลงพีเอสเอ็มเป็นโค้ด การแปลงทำโดยใช้เครื่องมือแปลงซึ่งทำงานตามกฎการแปลงที่กำหนด นักพัฒนาสามารถนำโค้ดที่แปลงได้ซึ่งอาจจะยังไม่สมบูรณ์มาทำการเพิ่มเติมให้กลายเป็นโค้ดโปรแกรมที่สมบูรณ์ต่อไป

ผู้วิจัยมีแนวคิดในการพัฒนาเว็บแอปพลิเคชันบนแพลตฟอร์มกูเกิลแอปเอนจินตามแบบเอ็มดีเอ โดยเสนอยูเอ็มแอลโปรไฟล์ [3] ซึ่งเป็นวิธีการระบุส่วนขยายให้กับยูเอ็มแอลสำหรับการสร้างแบบจำลองในโดเมนงานที่เฉพาะเจาะจง ซึ่งในที่นี้คือยูเอ็มแอลโปรไฟล์สำหรับเว็บแอปพลิเคชันบนแพลตฟอร์มกูเกิลแอปเอนจิน (UML profile for Web application on Google App Engine: WebGAE) สำหรับใช้สร้างแบบจำลองในระดับพีเอสเอ็มที่เฉพาะเจาะจงกับแพลตฟอร์มกูเกิลแอปเอนจิน แล้วจึงใช้กฎการแปลงที่ผู้วิจัยเสนอในการแปลงแบบจำลองไปเป็นโค้ดเทมเพลตของเว็บแอปพลิเคชันที่ใช้เทคโนโลยีจาวา สำหรับให้นักพัฒนาทำการเพิ่มเติมให้เป็นโค้ดที่สมบูรณ์ต่อไป ผู้วิจัยได้พัฒนาเครื่องมือสำหรับการแปลง รวมถึงนำเสนอกรณีศึกษาที่นำยูเอ็มแอลโปร

ไฟล์ที่เสนอไปประยุกต์ใช้งาน เพื่อประเมินประสิทธิภาพของการแปลง โดยพิจารณาสัดส่วนการแปลงโค้ดที่สร้างได้เทียบกับปริมาณโค้ดในแอปพลิเคชันที่นักพัฒนาทำการเพิ่มเติมจนสมบูรณ์แล้วเพื่อประเมินสัดส่วนการลดภาระของนักพัฒนาในการเขียนโค้ดเองทั้งหมด

ส่วนที่ 2 ของบทความนี้กล่าวถึงงานวิจัยที่เกี่ยวข้อง ส่วนที่ 3 กล่าวถึงขั้นตอนการพัฒนาเว็บแอปพลิเคชันบนกูเกิลแอปเอนจินแบบเอ็มดีเอและเครื่องมือสนับสนุน ส่วนที่ 4 กล่าวถึงกรณีศึกษาและผลการประเมิน และส่วนที่ 5 เป็นบทสรุป

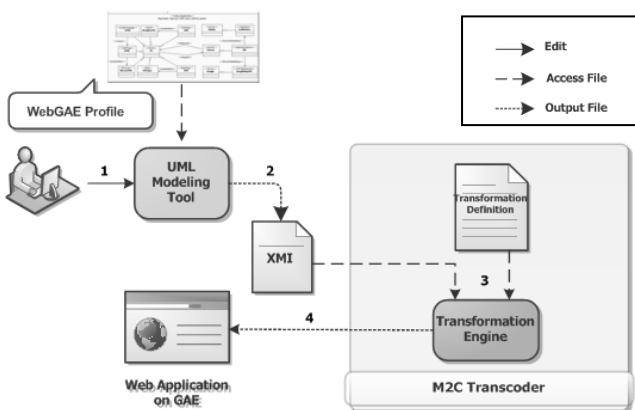
2. งานวิจัยที่เกี่ยวข้อง

Huang และคณะ [4] ได้เสนอการแปลงแผนภาพคลาสของยูเอ็มแอลไปเป็นโค้ดเทมเพลตของเว็บแอปพลิเคชันตามแบบรูป MVC (Model-View-Controller) โดยทำการขยายแบบจำลองแผนภาพคลาสของยูเอ็มแอลโดยกำหนดแม่พิมพ์ต้นแบบ (Stereotypes) ซึ่งแสดงถึงเอลิเมนต์เฉพาะสำหรับแบบจำลองของเว็บแอปพลิเคชัน และแม่พิมพ์ต้นแบบของความสัมพันธ์ (Association Stereotypes) ซึ่งแสดงถึงรูปแบบความสัมพันธ์ระหว่างเอลิเมนต์เฉพาะตามแบบจำลอง งานวิจัยของ Hsu [5] นำเสนอส่วนขยายของยูเอ็มแอลที่เรียกว่ายูเอ็มแอลโปรไฟล์สำหรับการสร้างแบบจำลองเว็บ 2.0 แมซฮัพ (Web2.0 MUML) เพื่อใช้ในการสร้างแบบจำลองในระดับพีไอเอ็มสำหรับการออกแบบเว็บแมซฮัพซึ่งเป็นการรวมเนื้อหาบนเว็บจากหลายแหล่งมานำเสนอบนหน้าเว็บเดียวกัน จากนั้นทำการแปลงแบบจำลองจากระดับพีไอเอ็มไปเป็นพีเอสเอ็มและโค้ดตามลำดับ โดยในยูเอ็มแอลโปรไฟล์จะกำหนดแม่พิมพ์ต้นแบบสำหรับสถาปัตยกรรมของ Application, Web Resource, Client และ Relationship ใ้งานวิจัยของ Kataria และ Yadav [6] เสนอวิธีการสร้างแบบจำลองของเว็บแอปพลิเคชันโดยสร้างส่วนขยายของยูเอ็มแอลจากมุมมองของส่วนประกอบต่าง ๆ ในเว็บแอปพลิเคชันซึ่งมีอยู่สอดคล้องกัน ไม่ว่าจะใช้เทคโนโลยีใดในการพัฒนาก็ตาม ทั้ง HTML, ASP, JSP, PHP, Servlet และ JavaBeans แม่พิมพ์ต้นแบบที่นำเสนอขึ้นครอบคลุมส่วนประกอบที่เป็นไคลเอนต์ เซิร์ฟเวอร์และความสัมพันธ์ภายในเว็บแอปพลิเคชัน

จากตัวอย่างงานวิจัยข้างต้น ผู้วิจัยได้แนวทางสำหรับการแปลงแบบจำลองเป็นโค้ดและตัวอย่างการสร้างยูเอ็มแอลโปรไฟล์สำหรับเว็บแอปพลิเคชัน แต่ยังไม่มียานใดที่เน้นการพัฒนาแบบเอ็มดีเอสำหรับเว็บแอปพลิเคชันที่สามารถทำงานบนแพลตฟอร์มกูเกิลแอปเอนจิน งานวิจัยของ Tsai และคณะ [7] นำเสนอวิธีการสร้างโค้ดสำหรับผู้ให้บริการซอฟต์แวร์บนคลาวด์ซึ่งบริการซอฟต์แวร์นั้นทำงานอยู่บนบริการแพลตฟอร์มบนคลาวด์ เช่น กูเกิลแอปเอนจิน โดยให้สามารถปรับแต่งเงื่อนไขเกี่ยวกับสภาพแวดล้อมของการทำงานของบริการซอฟต์แวร์สำหรับลูกค้าแต่ละราย แต่งานวิจัยดังกล่าวไม่ได้เน้นที่การพัฒนาเว็บแอปพลิเคชันและไม่ได้ใช้แนวทางของยูเอ็มแอลและเอ็มดีเอ

3. การพัฒนาเว็บแอปพลิเคชันบนกูเกิลแอปเอนจินแบบเอ็มดีเอ

ภาพรวมของการพัฒนาเว็บแอปพลิเคชันบนแพลตฟอร์มกูเกิลแอปเอนจินเป็นดังภาพที่ 1



ภาพที่ 1: ภาพรวมการพัฒนา

เริ่มต้นจากนักพัฒนาทำการออกแบบแผนภาพคลาสของระบบเว็บแอปพลิเคชันตามยูเอ็มแอลโปรไฟล์สำหรับเว็บแอปพลิเคชันบนแพลตฟอร์มกูเกิลแอปเอนจิน หรือ WebGAE โดยใช้เครื่องมือสร้างแบบจำลอง (Modeling Tool) จากนั้นนำแบบจำลองในระดับพีเอสเอ็มที่ได้ซึ่งอยู่ในรูปของไฟล์ XMI นำเข้าไปยังเครื่องมือการแปลง (Transformation Engine) ร่วมกับนิยามการแปลง (Transformation Definition) ซึ่งประกอบด้วยชุดของกฎการแปลง เพื่อใช้สำหรับดำเนินการ

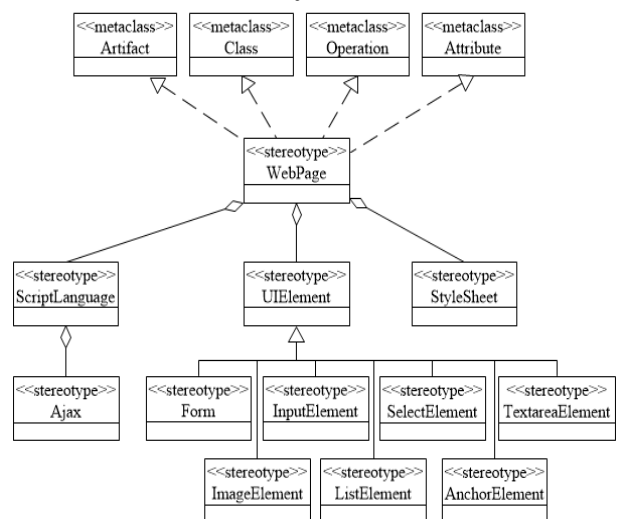
แปลงแบบจำลองพีเอสเอ็ม ผลลัพธ์ที่ได้จากเครื่องมือการแปลงคือ โค้ดเทมเพลตของเว็บแอปพลิเคชันที่ใช้เทคโนโลยีจาวาสำหรับทำงานบนแพลตฟอร์มกูเกิลแอปเอนจิน องค์ประกอบแต่ละส่วนอันได้แก่ 1) ยูเอ็มแอลโปรไฟล์ WebGAE 2) กฎการแปลงแบบจำลองเป็นโค้ด และ 3) เครื่องมือการแปลงแบบจำลองเป็นโค้ด มีรายละเอียดดังนี้

3.1 ยูเอ็มแอลโปรไฟล์ WebGAE

ผู้วิจัยทำการศึกษายูเอ็มแอลเมตาโมเดลเพื่อนำมาอ้างอิงในการสร้างส่วนขยายของยูเอ็มแอลในรูปแบบของยูเอ็มแอลโปรไฟล์สำหรับเว็บแอปพลิเคชันบนแพลตฟอร์มกูเกิลแอปเอนจิน หรือ WebGAE ในโปรไฟล์ประกอบด้วยแม่พิมพ์ต้นแบบ (Stereotype) ที่แสดงถึงเอลิเมนต์เฉพาะสำหรับแบบจำลองของเว็บแอปพลิเคชันบนกูเกิลแอปเอนจิน โดยแบ่งออกเป็นสามกลุ่มสถาปัตยกรรมดังนี้

3.1.1 สถาปัตยกรรมไคลเอนต์

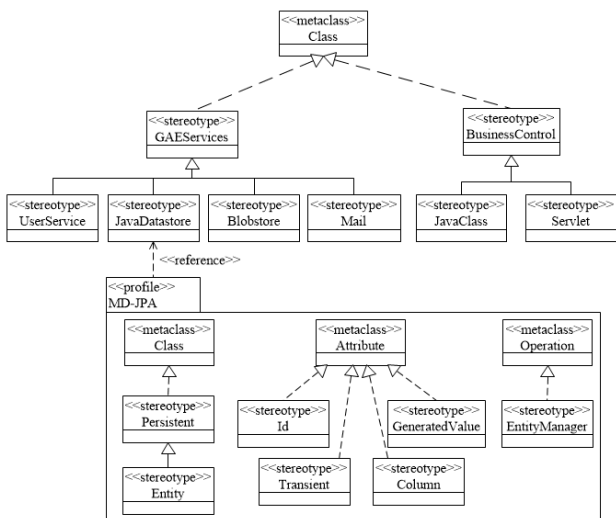
สถาปัตยกรรมไคลเอนต์อธิบายแบบจำลองการทำงานฝั่งไคลเอนต์ของเว็บแอปพลิเคชัน จากงานวิจัย [5] และ [6] นำเสนอการใช้งานยูเอ็มแอลโปรไฟล์โดยมีการกำหนดแม่พิมพ์ต้นแบบสำหรับส่วนประกอบของไคลเอนต์ จึงสามารถนำมาพัฒนาต่อสำหรับการพัฒนาเว็บแอปพลิเคชันในส่วนของแบบจำลองส่วนแสดงผล แม่พิมพ์ต้นแบบของไคลเอนต์เป็นส่วนขยายของเมตาคลาส Artifact, Class, Operation และ Attribute ในเมตาโมเดลของยูเอ็มแอล ดังภาพที่ 2



ภาพที่ 2: แม่พิมพ์ต้นแบบสำหรับสถาปัตยกรรมไคลเอนต์ (ปรับจาก [5], [6])

3.1.2 สถาปัตยกรรมเซิร์ฟเวอร์

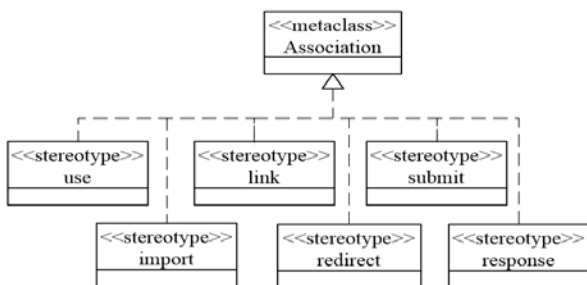
สถาปัตยกรรมเซิร์ฟเวอร์อธิบายการทำงานฝั่งเซิร์ฟเวอร์ของเว็บแอปพลิเคชัน งานวิจัยนี้ใช้กูเกิลแอปเอนจินเป็นเซิร์ฟเวอร์ ซึ่งสนับสนุนการพัฒนาด้วย Servlet, JSP และการจัดการข้อมูลผ่าน Java Persistence API (JPA) ผู้วิจัยได้นำโปรไฟล์ MD-JPA [8] มาใช้ร่วมด้วย ซึ่งการจัดการกับข้อมูลแบบไม่มีสคีมา (Schemaless) จะเก็บข้อมูลในรูปของ Entity นอกจากนี้กูเกิลแอปเอนจินยังมีบริการให้เรียกใช้ เช่น บริการ UserService, BlobStore และ Mail ภาพที่ 3 แสดงแม่พิมพ์ต้นแบบสำหรับฝั่งเซิร์ฟเวอร์ซึ่งเป็นส่วนขยายมาจากเมตาคลาสชื่อว่า Class ในเมตาโมเดลของยูเอ็มแอล



ภาพที่ 3: แม่พิมพ์ต้นแบบสำหรับสถาปัตยกรรมเซิร์ฟเวอร์ (ปรับจาก [6], [8])

3.1.3 สถาปัตยกรรมความสัมพันธ์

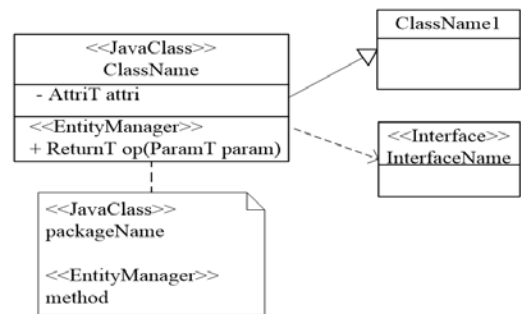
สถาปัตยกรรมความสัมพันธ์อธิบายความสัมพันธ์ระหว่างแต่ละคลาสของเว็บแอปพลิเคชัน ซึ่งแม่พิมพ์ต้นแบบของความสัมพันธ์เป็นส่วนขยายมาจากเมตาคลาสชื่อว่า Association ในเมตาโมเดลของยูเอ็มแอล ดังภาพที่ 4



ภาพที่ 4: แม่พิมพ์ต้นแบบสำหรับสถาปัตยกรรมความสัมพันธ์ (ปรับจาก [6], [8])

3.2 กฎการแปลงแบบจำลองเป็นโค้ด

กฎการแปลงสำหรับแปลงแบบจำลองซึ่งสร้างจากแม่พิมพ์ต้นแบบในโปรไฟล์ WebGAE ไปเป็นโค้ด สามารถอธิบายในรูปแบบไวยากรณ์อีบีเอ็นเอฟ (EBNF – Extend Backus Naur Form) เนื่องด้วยข้อจำกัดด้านเนื้อหาที่ ในที่นี้จะยกตัวอย่างกฎการแปลงสำหรับแบบจำลองที่ใช้แม่พิมพ์ต้นแบบของ JavaClass ดังในภาพที่ 5 แม่พิมพ์ต้นแบบจะมีนิยามของป้ายระบุ (Tag Definition) ดังแสดงด้วยสัญลักษณ์ Comment ซึ่งช่วยกำหนดคุณลักษณะเฉพาะของแม่พิมพ์ต้นแบบและมีการกำหนดค่า (Tagged Value) ให้เมื่อนำแม่พิมพ์ต้นแบบไปใช้ในการออกแบบ



ภาพที่ 5: แบบจำลองที่ออกแบบโดยใช้แม่พิมพ์ต้นแบบของ JavaClass

ขั้นตอนการแปลงแบบจำลองที่ออกแบบโดยใช้แม่พิมพ์ต้นแบบของ JavaClass ไปเป็นโค้ด มีดังนี้

1. กฎการแปลง <<JavaClass>> ไปเป็นโค้ดแสดงดังภาพที่ 6 ชื่อคลาส ClassName ในภาพที่ 5 จะถูกแปลงเป็นชื่อคลาส <<JavaClass>>ClassName ในชุดคำสั่งภาษาจาวา ซึ่งมีการระบุระดับการเข้าถึงเป็น public
2. กฎการแปลงแอตทริบิวต์ใน <<JavaClass>> ไปเป็นโค้ดแสดงดังภาพที่ 7 ชื่อแอตทริบิวต์ attri และประเภทข้อมูล AttriT ในภาพที่ 5 จะถูกแปลงเป็นแอตทริบิวต์ ATTRIBUTENAME_{attri} ที่มีประเภทข้อมูล TYPE_{AttriT} ในคลาสในชุดคำสั่งภาษาจาวา

```

<<CLASS-DECLARATION>> ::=
"package" <<JavaClass>>packageName ";"
{ <<IMPORT-STATEMENT>> }
"public class" <<JavaClass>>ClassName
[ "extends" ClassName1 ]
[ "implements" <<Interface>>InterfaceName
{ ";" <<Interface>>InterfaceName } ]
"{" { <<CLASS-BODY>>JavaClass } "}"

<<CLASS-BODY>>JavaClass ::=
( <<ATTRIBUTE>>JavaClass | <<METHOD>>JavaClass )
    
```

ภาพที่ 6: กฎการแปลงสำหรับ <<JavaClass>>

```
<ATTRIBUTEJavaClass> ::=
    { <MODIFIER> } TYPE_AttriT ATTRIBUTEName_Attri
    { “,” ATTRIBUTEName_Attri } “;”
```

ภาพที่ 7: กฎการแปลงสำหรับแอตทริบิวต์ใน <<JavaClass>>

3. กฎการแปลงเมทอดใน <<JavaClass>> ไปเป็นโค้ด แสดงดังภาพที่ 8 ชื่อเมทอด op ในภาพที่ 5 จะถูกแปลงเป็นชื่อ เมทอด METHODNAME_{op} ชื่อพารามิเตอร์ param ที่มี ประเภท ข้อมูล ParamT จะ ถูก แปลง เป็น พารามิเตอร์ PARAMNAME_{param} ที่มีประเภท TYPE_{paramT} และประเภทข้อมูลตอบกลับ ReturnT จะถูกแปลงเป็นประเภท ข้อมูล TYPE_{ReturnT}

```
<METHODJavaClass> ::=
    { <MODIFIER> } TYPE_ReturnT METHODNAME_op
    “(“ [ <PARAMETER-LIST> ] “)”
    “{”
        [ <ENTITYMANAGER> ]
        [ <RETURN-VALUEJavaClass> ]
    “}”
<PARAMETER-LIST> ::=
    TYPE_ParamT PARAMNAME_param
    { “,” TYPE_ParamT PARAMNAME_param }
<ENTITYMANAGER> ::=
    “EntityManager em = EMF.get().createEntityManager();”
    “try {”
        “em.” <<EntityManager>>_method “(/*Insert code here*/);”
    “} catch (Exception ex) {”
        “ex.printStackTrace();”
    “} finally {”
        “em.close();”
    “}”
<RETURN-VALUEJavaClass> ::= “return “/*Insert code here*/;”
```

ภาพที่ 8: กฎการแปลงสำหรับเมทอดใน <<JavaClass>>

นอกจากนี้หากเมทอดมีการระบุด้วยแม่พิมพ์ต้นแบบ <<EntityManager>> ดังเช่นภาพที่ 5 แสดงว่าในการทำงานของเมทอดจะมีการจัดการข้อมูลซึ่งเป็น Entity ใน DataStore ของกูเกิลแอปเอนจิน วิธีการจัดการข้อมูลได้แก่ เมทอดเพิ่ม (persist) แก้ไข (merge) และลบ (remove) ข้อมูล Entity นักพัฒนาสามารถระบุเพิ่มเติมลงในโค้ดเทมเพลตในภายหลังได้ว่า จะจัดการข้อมูล Entity ใด ส่วนจะจัดการวิธีใดนั้น นักพัฒนาสามารถระบุเป็นค่า Tagged Value ให้กับ Tag Definition method ของ <<EntityManager>> ได้ตอนที่ ออกแบบ ค่า Tagged Value นี้ จะถูกแปลง เป็น <<EntityManager>>_method ในชุดคำสั่งภาษาจาวา

3.3 เครื่องมือการแปลงแบบจำลองเป็นโค้ด

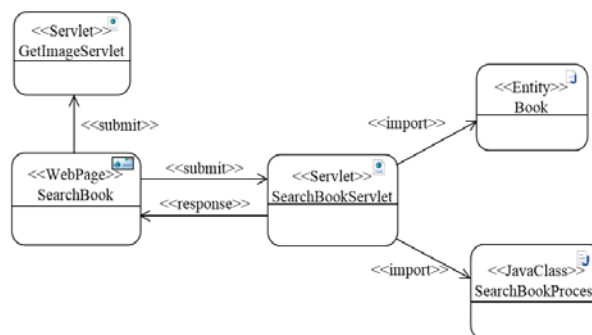
เครื่องมือการแปลงแบบจำลองเป็นโค้ดนั้นพัฒนาขึ้นในรูปแบบของจาวาแอปพลิเคชัน โดยสามารถรองรับการแปลงแบบจำลองที่กำหนดขึ้น ฟังก์ชันการทำงานของเครื่องมือประกอบด้วยการนำเข้าแบบจำลองพีเอสเอ็มในรูปแบบไฟล์ XMI ซึ่งได้มาจากการใช้เครื่องมืออื่น เช่น Enterprise Architect ในการสร้างแบบจำลองตามโปรไฟล์ WebGAE จากนั้นกำหนดไลบรารีสำหรับกูเกิลแอปเอนจินให้กับเครื่องมือ และเลือกไดรเวทอรี่สำหรับจัดเก็บโค้ดหลังจากแปลงเสร็จ เมื่อการแปลงเสร็จเรียบร้อยแล้ว นักพัฒนาสามารถนำโค้ดเทมเพลตที่ได้ไปเพิ่มเติมให้สมบูรณ์ต่อไป

4. การประยุกต์และประเมินผล

ผู้วิจัยประยุกต์วิธีที่เสนอเข้ากับกรณีศึกษาเพื่อใช้ในการประมวลผล

4.1 การพัฒนาเว็บแอปพลิเคชันของกรณีศึกษา

ผู้วิจัยใช้ระบบสั่งซื้อหนังสือออนไลน์เป็นกรณีศึกษา ระบบมีฟังก์ชันงานหลัก 5 ฟังก์ชัน ได้แก่ ล็อกอิน (User Login) ค้นหารายการหนังสือ (Search Book List) ตะกร้าสินค้า (Shopping Cart) สั่งซื้อสินค้า (Payment Order) และเพิ่มหนังสือใหม่ (Add New Book) ภาพที่ 9 เป็นตัวอย่างแผนภาพคลาสของระบบในส่วนที่เกี่ยวกับฟังก์ชันการค้นหารายการหนังสือ โดยมีการใช้แม่พิมพ์ต้นแบบของโปรไฟล์ WebGAE เพียงบางส่วน เมื่อออกแบบแล้ว ได้นำแบบจำลองไปแปลงเป็นโค้ดโดยใช้เครื่องมือการแปลง จากนั้นให้นักพัฒนาเพิ่มโค้ดให้เป็นแอปพลิเคชันที่สมบูรณ์และทำงานได้บนกูเกิลแอปเอนจิน



ภาพที่ 9: ตัวอย่างแผนภาพคลาสของฟังก์ชันการค้นหารายการหนังสือ

4.2 การประเมินผล

ผู้วิจัยทำการประเมินประสิทธิภาพของการแปลงแบบจำลองไปเป็นโค้ด แต่เนื่องจากโค้ดที่เครื่องมือแปลงได้ อาจเป็นโค้ดที่ไม่ได้แสดงฟังก์ชันงานที่สมบูรณ์ และนักพัฒนา ยังต้องเพิ่มโค้ดเพื่อให้ทำงานได้ตามฟังก์ชันที่ต้องการ ผู้วิจัยจึงทำการวัดสัดส่วนโค้ดที่แปลงได้โดยตรงโดยนับจำนวนบรรทัด (Line of Code: LOC) ของโค้ดที่สร้างได้สำหรับกรณีศึกษา เทียบกับปริมาณโค้ดในแอปพลิเคชันที่นักพัฒนาทำการเพิ่มเติม จนสมบูรณ์แล้ว [4] ดังผลการประเมินในตารางที่ 1

ตารางที่ 1: ผลการประเมินการแปลงแบบจำลองเป็นโค้ด

	LOC ของโค้ด เทมเพลต	LOC ของโค้ดที่ เสร็จสมบูรณ์	อัตราการ แปลง
Use Case 1: User Login (MainPageServlet.java, RegisterProcess.java, Customer.java)	114	144	79%
Use case 2: Search Book List (searchBook.jsp, SearchBookServlet.java, GetImageServlet.java, SearchBookProcess.java)	133	209	64%
Use case 3: Shopping Cart (viewCart.jsp, ShoppingCartServlet.java, ShoppingCartProcess.java)	137	305	45%
Use case 4: Payment Order (payment.jsp, paymentSuccess.jsp, OrderServlet.java, OrderProcess.java, Order.java, OrderDetail.java)	259	374	69%
Use case 5: Add New Book (addBook.jsp, AddBookServlet.java, ManageBookProcess.java, Book.java)	123	162	76%

จากผลการประเมินในตารางที่ 1 พบว่าสัดส่วนการแปลง (Transformation Ratio) อยู่ที่ 45 – 79% ในส่วนของโค้ดที่สามารถกำหนดด้วยแบบจำลองโครงสร้างเชิงสถิติ ประกอบด้วยแอตทริบิวต์ เมทอด และความสัมพันธ์ระหว่างคลาส และไม่ได้มุ่งเน้นส่วนตรรกะของโปรแกรมอย่างเช่น คลาส Entity หน้าเว็บ จะมีสัดส่วนการแปลงที่สูง แต่สำหรับโค้ดที่ต้องการการกำหนดรายละเอียดตรรกะของโปรแกรม จะมีอัตราการแปลงที่ต่ำลง เนื่องจากยังต้องอาศัยนักพัฒนาในการเพิ่มเติมโค้ดให้สมบูรณ์

5. บทสรุป

นักพัฒนาเว็บแอปพลิเคชันบนแพลตฟอร์มกูเกิลแอปเอนจินจะได้รับความสะดวกจากการใช้เครื่องมือที่สนับสนุนการ

แปลงแบบจำลองที่สร้างด้วยโปรไฟล์ WebGAE และใช้กฎการแปลงที่ผู้วิจัยเสนอในการพัฒนา เนื่องจากเครื่องมือสามารถสร้างโค้ดเทมเพลตให้อย่างอัตโนมัติในสัดส่วนที่ค่อนข้างสูงในอนาคตสามารถพัฒนางานวิจัยต่อให้รองรับการแปลงแผนภาพยูเอ็มแอลเชิงกิจกรรม เช่น แผนภาพลำดับ ไปเป็นโค้ด เพื่อให้ช่วยเพิ่มสัดส่วนการแปลงให้สูงขึ้น อีกทั้งยังสามารถปรับปรุงกฎการแปลงเพื่อให้โค้ดที่แปลงได้เป็นไปตามแบบรูปการออกแบบ (Design Patterns) ที่เหมาะสมได้

เอกสารอ้างอิง

- [1] Google, *Google App Engine*, <https://developers.google.com/appengine/>, Last access: 28 February 2014.
- [2] A. Kleppe, J. Warmer, and W. Bast, *MDA Explained: The Model Driven Architecture: Practice and Promise*, Pearson Education, Inc, 2003.
- [3] K. Hamilton and R. Miles, *Learning UML 2.0*, O'Reilly, 2006.
- [4] Y.-C. Huang, C.-P. Chu, Z.-A. Lin, and M. Matuschek, "Transformation from Web PSM to code," *Proc. 15th Int. Conf. Distributed Multimedia Systems (DMS 2009)*, 12-14 Sep 2009, pp. 16-19.
- [5] I.-C. Hsu, "Visual modeling for Web 2.0 applications using model driven architecture approach," *Simulation Modelling Practice and Theory*, Vol. 31, pp. 63-76, 2013.
- [6] M. Kataria and R. Yadav, "A component-centric UML based approach for modeling the architecture of web applications," *Int. J. Recent Research and Review*, Vol. V, pp. 22-27, Mar 2013.
- [7] W.-T. Tsai, W. Li, B. Esmaili, and W. Wu, "Model-driven tenant development for PaaS-based SaaS," *Proc. 4th IEEE Int. Conf. Cloud Computing Technology and Science (CloudCom 2012)*, 3-6 Dec 2012, pp. 821-826.
- [8] A. Torres, R. Galante, and M.S. Pimenta, "Towards a UML profile for model-driven object-relational mapping," *Proc. XXIII Brazilian Symposium on Software Engineering*, 2009, pp. 94-103.

**ระบบการเผยแพร่บทความเพื่อสนับสนุนการทำงาน
ของนักการตลาดบนอินเทอร์เน็ตแบบเน้นจำนวนบทความ
Automatic System of Post Master Supported Affiliate Marketer
by Number of Post Method**

กิตตินันท์ น้อยมณี (Kittinan Noimane)¹, รัชชาติ ชูเกียรติจิกร (Natchasit Chukiathajorn)²

ยุพดี หัตถสิน (Upady Hatthasin)³ และ อนุพงษ์ ไพโรจน์ (Anupong Pairote)⁴

สาขาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนาภาคพายัพเชียงใหม่

128 ถนนห้วยแก้ว ตำบลช้างเผือก อำเภอเมืองเชียงใหม่ จังหวัดเชียงใหม่ ประเทศไทย

mr.kittinan@rmutl.ac.th¹, gohgoh@rmutl.ac.th², uht@rmutl.ac.th³, anupong@rmutl.ac.th⁴

บทคัดย่อ

ระบบเผยแพร่บทความอัตโนมัติสร้างขึ้นเพื่อสนับสนุนการทำงานของนักการตลาดบนอินเทอร์เน็ตเป็นหลักที่มุ่งเน้นจำนวนบทความเพื่อขยายผลต่อการซื้อขายในอนาคต เพราะบทความเป็นส่วนสำคัญในการเพิ่มแรงจูงใจในการซื้อหรือขายสินค้าของมนุษย์งานวิจัยนี้จึงได้เน้นการสร้างบทความเป็นหลัก โดยจะครอบคลุมจนถึงการติดอันดับบนระบบของ Search Engine ซึ่งใช้เทคโนโลยี Smarty Template Intelligent ที่ช่วยในการร่างโครงสร้างของเว็บไซต์ให้สนับสนุนกับการเก็บข้อมูลของ Search Engine โดยหุ่นยนต์ประจำ Search Engine แต่ละค่าย โดยเป็นการปรับแต่งตามหลักการของ Search Engine Optimization ประเภท On-page ส่วนแบบ Off-page จะทำการปรับแต่งโดยใช้ระบบการสร้างสะพานให้กับหุ่นยนต์ประจำ Search Engine โดยใช้กระบวนการ Link Wheel

คำสำคัญ: การเผยแพร่บทความอัตโนมัติ นักการตลาดบนอินเทอร์เน็ต

Abstract

Post Master is automatic posting article that support affiliate marketer. It can be increased the value of article which have high quality and impressive in the internet. This research will explain the method of post master that provide Smarty template intelligent, structure's supported by search engine robot, very helpful for Search Engine Optimization(SEO) on on-page and off-page type. On-page type, smarty template intelligent have many facility

that provided and very useful. Off-page type, researcher use link wheel algorithm for analysis this problem that build provide search engine robot across the link bridge, increase index ratio from search engine.

Keyword: Post Master, Affiliate Marketer

1. บทนำ

การตลาดออนไลน์หรือการตลาดบนระบบเครือข่ายอินเทอร์เน็ตหมายถึงกิจกรรมหรือเทคนิคบางอย่างที่สามารถทำให้ธุรกิจนั้นๆ ประสบความสำเร็จหรือสามารถไปถึงเป้าหมายที่ตั้งไว้ได้ ซึ่งจะมีหลักการคล้ายกับการตลาดโดยทั่วไปแต่จะแตกต่างกันส่วนของวิธีการให้ได้มาในการทำการตลาดนั้นๆ และมีหลายรูปแบบ เช่น การโฆษณาสินค้าผ่านระบบเครือข่ายอินเทอร์เน็ต[1] การสร้างคำโฆษณาผ่านระบบเครือข่ายอินเทอร์เน็ต[3-4] การเปรียบเทียบคุณสมบัติและประโยชน์ของสินค้าในแต่ละประเภทผ่านระบบเครือข่ายอินเทอร์เน็ต หรือการสร้างบทความเพื่อส่งเสริมการขายสำหรับสินค้าชนิดต่างๆ [5-8] ดังนั้นการทำการตลาดทุกครั้งจะต้องทำการวิจัยสินค้าพร้อมทั้งวิจัยตลาดก่อนทุกครั้ง[2, 6, 8-9]

2. ระบบการเผยแพร่บทความอัตโนมัติ

2.1 อัลกอริทึมของ Search Engine

Search Engine ของแต่ละบริษัทจะเป็นผู้ให้บริการในการค้นหา เพื่อให้ผู้ใช้บริการบรรลุเป้าหมาย และจัดอันดับของเนื้อหาของเว็บต่างๆ ทั่วโลก ซึ่งขั้นตอนและกระบวนการคิด

ของ Search Engine ในแต่ละบริษัทจะมีลักษณะที่คล้ายคลึงกัน โดยจะทำการสร้างระบบเพื่อทำการตรวจสอบและจัดอันดับเว็บไซต์ทั้งหมดของเซิร์ฟเวอร์และผู้ให้บริการเซิร์ฟเวอร์ฟรี ระบบดังกล่าวคือ Search Engine Spider ซึ่งมีหน้าที่ในการจัดอันดับและวิเคราะห์โครงสร้างพื้นฐานของแต่ละเว็บเพจที่มีอัตราการเติบโตในลักษณะรายวัน ซึ่งหลังจากทำการเก็บข้อมูล Search Engine Spider จะทำการนำข้อมูลดังกล่าวมาทำการประมวลผลที่ Search Engine โดยจะทำการวิเคราะห์และจัดอันดับคะแนนของเว็บไซต์ที่ได้ทำการเก็บข้อมูลมา ดังนั้นลักษณะการสร้างบทความแบบเน้นจำนวนบทความจึงต้องอาศัยหลักการของ Link Wheel เพื่อการเพิ่มช่องทางและโอกาสในการค้นพบเว็บไซต์ของนักการตลาดบนอินเทอร์เน็ต ซึ่งอธิบายได้ดังภาพที่ 1 โดยสมมติฐานที่ว่า Search Engine Spider จะมีความถี่ในการเก็บข้อมูลจากเว็บบล็อกสูง ดังนั้นการส่งเสริมการขายโดยใช้คำโฆษณาผ่านเว็บบล็อกจึงจะเพิ่มโอกาสในการเกิดกระบวนการซื้อขายสินค้าเพื่อให้เกิดประโยชน์ต่อนักการตลาดบนเครือข่ายอินเทอร์เน็ตมากยิ่งขึ้น ซึ่งเป็นลักษณะโครงสร้างการทำ SEO ประเภท Off-page



ภาพที่ 1: Link Wheel เชื่อมระหว่าง Web 2.0 และเว็บไซต์หลัก

2.2 Search Engine Optimization

Search Engine Optimization (SEO) คือมาตรฐานการจัดการโครงสร้างของเว็บไซต์เพื่ออำนวยความสะดวกให้กับ Search Engine Spider ในการเก็บข้อมูล[2] แบ่งออกเป็น 2 ประเภทคือ

ประเภทแรก การทำ SEO ในลักษณะ On-page เป็นลักษณะของการปรับโครงสร้างของเว็บไซต์เพื่ออำนวยความสะดวกในการเก็บข้อมูลจาก Search Engine Spider เช่นการใช้งาน Tag เพื่อทำการแบ่งส่วนของเนื้อหา การจัดตำแหน่งของหัวข้อ ลำดับของหัวข้อ ลักษณะของตัวอักษร ความหนาแน่นของคำหลักที่เกี่ยวข้องบทความ อัตราเร็วในการเก็บข้อมูลจะขึ้นอยู่กับโครงสร้างที่ไม่ซับซ้อน **ประเภทที่สอง** การทำ SEO ในลักษณะ Off-page เป็นลักษณะของความน่าเชื่อถือของเว็บไซต์ โดยการอ้างอิงจากเว็บไซต์ที่น่าเชื่อถือและมีอัตราการเกิดกิจกรรมภายในเว็บไซต์สูง โดย Search Engine Spider จะทำการวิเคราะห์จากลิงก์ภายนอก (Back Link: BL) ที่เกิดขึ้นภายนอกเว็บไซต์หลัก เช่นลิงค์ของเว็บไซต์หลักที่อยู่ในเว็บบล็อกที่มีการปรับปรุงและเพิ่มเนื้อหาอย่างสม่ำเสมอ หากอัตราของกิจกรรมต่างๆ ในเว็บบล็อกสูง Search Engine Spider จะลงความเห็นให้เว็บไซต์หลักมีความเชื่อถือเช่นกัน

เพียงแต่การปรับโครงสร้างของเว็บไซต์ทั้งสองลักษณะจำเป็นต้องใช้ระยะเวลาที่มาก หากอัตราการเติบโตของบทความสูง ย่อมส่งผลให้ระยะเวลาในการปรับโครงสร้างของเว็บไซต์สูงด้วยเช่นกัน ผู้วิจัยจึงใช้เทคโนโลยีของ Smarty[4] ที่เป็น Template Intelligent เข้ามาช่วยระยะเวลาของขั้นตอนดังกล่าวซึ่งแสดงผลดังภาพที่ 2

On New Drum Set Black 5-Piece Complete Full Size with Cymbal Mon Percussion BATTLEBK (Black Friday and Cyber Monday)

ack 5-Piece Complete Full Size with Cymbals Stands Stool Sticks ? At this time Most people who's got the best price on it recommend to buy it from there - You might also get Fast & Free shipping with Special offer.

Full Size with Cymbals Stands Stool Sticks Save up to 33%



Last Price : 5640.06

See Discount & Compare Price

5-Piece Complete Full Size with Cymbals Stands Stool Sticks **Specifics**

Everything Included
Stool Stands all Included
Inexpensive Site Set - Nothing Else To Buy
Driven Pedals
also - Best Price

ภาพที่ 2: การปรับโครงสร้าง SEO ในลักษณะ On-page โดยใช้ Smarty Template Intelligent Template

2.3 การเผยแพร่บทความบนระบบเครือข่ายอินเทอร์เน็ต

หน้าที่หลักของนักการตลาดบนระบบเครือข่ายอินเทอร์เน็ตคือเพิ่มอัตราการเข้าชมเว็บไซต์ที่มีคำสั่งเสริมการขายของสินค้าที่ได้ทำการวิจัยทางการตลาดผ่านบทความระดับคุณภาพ ดังนั้นจำนวนของบทความคุณภาพจึงมีความสอดคล้องกับอัตราเร็วในการจัดอันดับของ Search Engine Spider จึงเป็นเหตุผลที่ต้องทำการปรับโครงสร้างเว็บไซต์ให้เป็นมาตรฐานเพื่อเพิ่มอัตราเร่งในการจัดอันดับ พร้อมทั้งการเผยแพร่ผ่านระบบเว็บบล็อกที่ได้รับความนิยมเพื่อการกระตุ้นให้ Search Engine Spider เชื่อมโยงความสำคัญของเว็บไซต์หลักด้วยผู้วิจัยจึงเน้นไปยังเครือข่ายต่างๆ เช่น Facebook Wordpress Hatena Blog และ Blogger เพราะเป็นเครือข่ายเว็บไซต์ที่ได้รับความนิยมสูงโดยการจัดอันดับจาก Search Engine ให้มีค่า Page Rank อยู่ที่ 9 ซึ่งเป็นขั้นตอนวิธีในการวิเคราะห์เว็บลิงค์ตามทฤษฎีเครือข่ายที่เป็นพื้นฐานในตัว Search Engine

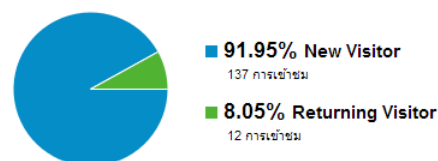
3. ส่วนการจัดการและการบริหารภายในระบบเผยแพร่บทความอัตโนมัติ

สำหรับแหล่งสินค้าที่นักการตลาดบนระบบเครือข่ายอินเทอร์เน็ตให้ความสนใจในปัจจุบันคือ Amazon.com เพราะมีบริการในการเรียกใช้คำสั่งในการดึงข้อมูลของสินค้าจากเว็บไซต์ได้ เมื่อทำการดึงข้อมูลจากระบบของ Amazon ระบบก็สามารถทำการจัดโครงสร้างต่างๆ ของเว็บเพจที่ใช้ในการสนับสนุนการขาย (Landing Page) เพื่อให้สอดคล้องกับโครงสร้างของ SEO ในลักษณะ On-page พร้อมทั้งเก็บลิงค์ของบทความพร้อมคำสำคัญของสินค้าแยกเพื่อเตรียมการทำ SEO ในลักษณะ Off-page หลังจากนั้นจะเข้าสู่กระบวนการนำข้อมูลที่ได้จาก Amazon ผสานเข้ากับโครงสร้างที่ได้ออกแบบไว้ด้วย

Smarty Template Intelligent พร้อมทั้งทำการเผยแพร่โครงสร้างดังกล่าวสู่เว็บไซต์หลักอันได้แก่ Facebook Wordpress Blogger และ Hatena Blog ได้โดยทันที โดยการเผยแพร่ผ่าน Facebook จะใช้วิธีการเชื่อมต่อผ่าน API และทาง Wordpress กับ Blogger จะเชื่อมต่อโดยผ่านการส่งชุดคำสั่ง XML และ Hatena Blog จะเชื่อมต่อโดยการส่งอีเมล ด้วยขั้นตอนวิธีต่างๆ ทำให้บทความที่ทำการจัดโครงสร้างไว้สามารถเผยแพร่ได้ทันทีโดยปริมาณในการเผยแพร่จะขึ้นอยู่กับจำนวนสินค้าที่ต้องการนำมาส่งเสริมการขาย

4. การอภิปรายผล

ในการทดสอบระบบ ผู้วิจัยได้ทำการติดตั้งรหัสเพื่อทำการติดตามผลของเว็บไซต์หลักที่ได้ทำการจัดโครงสร้างต่างๆ ซึ่งเครื่องมือที่นำมาวิเคราะห์คือ Google Analytic ทำให้ได้ผลลัพธ์คือ มีจำนวนผู้เข้าชมใหม่อย่างสม่ำเสมอในอัตรา 91.95% ด้วยการเข้าชมเป็นจำนวนทั้งหมด 194 หน้า มีจำนวนการเข้าชมที่ไม่ซ้ำกันคือ 138 คน โดยมีอัตราในการเข้าชมโดยเฉลี่ยคิดเป็น 42 วินาทีต่อหนึ่งเว็บเพจ ด้วยเนื้อหาที่สร้างเพื่อเน้นการนำพาผู้เข้าชมไปยังหน้าของสินค้าหลักไปยังเว็บ Amazon หรือว่าเป็นตัวเลขที่น่าพึงพอใจ สามารถอธิบายได้ดังภาพที่ 3 ซึ่งเปรียบเทียบกับภาพที่ 4 ที่แสดงจำนวนผู้เข้าชมเว็บไซต์ที่ผ่านกระบวนการซื้อขายสินค้า ซึ่งเกิดกระบวนการชำระสินค้าทั้งสิ้น 23 ชิ้น คิดเป็นราคาโดยรวม \$1,135.42 โดยนักการตลาดจะได้รับส่วนแบ่งทางการตลาดทั้งสิ้น \$60.05 ซึ่งสามารถทำการคำนวณอัตราความสำเร็จในการแปรเปลี่ยนจากผู้เข้าชมเว็บไซต์เป็นลูกค้าได้จาก อัตราส่วนของผู้ที่เป็ลูกค้าต่ออัตราส่วนของผู้ที่เข้าชมเว็บไซต์ทั้งหมดคือ 16.66%



ภาพที่ 3: สถิติการเข้าชมเว็บไซต์หลักที่ใช้ระบบเผยแพร่บทความอัตโนมัติโดยใช้ Google Analytic

August 8, 2012 to January 7, 2013			
	Items Shipped	Revenue	Advertising Fees
Total Amazon.com Items Shipped	12	\$447.83	\$20.76
Total Third Party Items Shipped 	11	\$687.59	\$39.29
Total Items Shipped	23	\$1,135.42	\$60.05
Total Items Returned	0	\$0.00	\$0.00
Total Refunds	0	\$0.00	\$0.00
TOTAL ADVERTISING FEES	23	\$1,135.42	\$60.05

ภาพที่ 4: สถิติจำนวนผู้เยี่ยมชมที่ผ่านกระบวนการซื้อขายสินค้าจาก Amazon ตั้งแต่วันที่ 8 สิงหาคม พ.ศ.2555 จนถึงวันที่ 7 มกราคม พ.ศ. 2556

5. สรุปผลการวิจัย

ด้วยอัตราส่วนของผู้ที่เป็นลูกค้าต่ออัตราส่วนของผู้ที่เยี่ยมชมเว็บไซต์ทั้งหมดคือ 16.66% หากจำนวนผู้เยี่ยมชมเว็บไซต์หลักมีปริมาณที่มากขึ้น จะส่งผลให้เกิดอัตราส่วนของผู้ที่เป็นลูกค้ามากขึ้นตามกัน ดังนั้นความเร็วในการทำอันดับของเว็บไซต์จึงมีอิทธิพลต่อนักการตลาดบนระบบเครือข่ายอินเทอร์เน็ตเป็นอย่างมาก และการเพิ่มจำนวนผู้เยี่ยมชมรายใหม่ก็ขึ้นอยู่กับจำนวนและระยะเวลาในการเผยแพร่บทความสู่เว็บบล็อกด้วยเช่นกัน

เอกสารอ้างอิง

- [1] Miodrag Ivkovic, Dusanka Milanov, "Affiliate Internet Marketing," International Conference on Education and Management Technology (ICEMT 2010), 978-1-4244-816-2/10, IEEE 2010.
- [2] HO Li-Hsing, Huang Jui-Chen, "The Application of Search Engine Optimization for Internet Marketing: An Example of the Model Websites," 978-1-4244-5586-7/10, IEEE 2010.
- [3] Kei-Shao Chen, Ming-Huei Chen, "EC 2.0: Can You Get Profit by Writing Blog? An Empirical Study in Google Adsense," 978-1-890843-21-0/10, IEEE 2010.
- [4] Salvatore Parise, Patricia J. Guinan, "Marketing Using Web 2.0," the 41st Hawaii International Conference on System Sciences, 2008, 1530-1605/08, IEEE 2008.
- [5] SUN Hua-Mei, LI YI-jun, HUANG Ti-Yun, "Research on Internet Marketing Strategies of Enterprises in China," School of Management, Harbin Institute of Technology, P.R. China, 150001, IEEE.
- [6] Cheng Congxi, Gen Pengfei, "Internet Marketing and Innovation Strategies: A Study of Marketing China's Agricultural Products," 978-1-4244-4639-1/09, IEEE 2009.
- [7] Dimitar Jovevski, Edvard Tijan, Petra Karanikic, "Internet Marketing Strategies and ICT as a Common Ground for Business Development," MIPRO 2010, May 24-28, 2010, Opatija, Croatia, IEEE 2010.
- [8] Wei Zhang, Jingdong Chen, Wei Han, "The Internet Marketing Strategies Theory and Empirical Research Based on Product Value," 978-1-4244-5265-1/10, IEEE 2010.
- [9] Meng Jiajia, Dong Dahai, "Determinants and Effect of Internet Marketing in E-Commerce: An Empirical Investigation from China," International Conference on E-Business and E-Government 2010, 978-0-7695-3997-3/10, IEEE 2010.

ตัวแบบสำหรับการเรียงลำดับผลลัพธ์การค้นคืนในระบบค้นคืนบทความวิจัยโดยใช้ข้อมูลทางบรรณานุกรม

A Model for Ranking Search Results in a Research Paper Search Engine Using Bibliographic Information

ขวัญเรือน โสอุบล (Khwanruan So-Ubol)¹ และ วรสิทธิ์ ชูชัยวัฒนา (Worasit Choochaiwattana)²

^{1,2}สาขาวิชาวิศวกรรมเว็บ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยธุรกิจบัณฑิตย์
ksoubol@gmail.com¹, worasit.cha@dpu.ac.th²

บทคัดย่อ

Query Dependent Ranking หรือ *Similarity Ranking* เป็นเทคนิคสำหรับเรียงลำดับผลลัพธ์การค้นคืนโดยการเปรียบเทียบคำค้นและดัชนีของเอกสาร ซึ่งไม่ได้พิจารณาถึงปัจจัยอื่นๆ เช่น คุณภาพของเอกสาร ในขณะที่ *Query Independent Ranking* หรือ *Static Ranking* เป็นเทคนิคที่สำคัญอีกเทคนิคหนึ่ง สำหรับการเรียงลำดับผลลัพธ์การค้นคืน โดยพิจารณาคุณภาพของเอกสารเป็นหลัก ในงานวิจัยนี้เสนอตัวแบบสำหรับเรียงลำดับผลลัพธ์การค้นคืนในระบบค้นคืนบทความวิจัย ที่มีการผสมผสานระหว่าง *Query Dependent Ranking* และ *Query Independent Ranking* โดยจะนำเอาข้อมูลทางบรรณานุกรม มาใช้สำหรับการประเมินคุณภาพของบทความวิจัย ซึ่งจะเป็นการใช้ *Similarity Feature* ประกอบด้วย ชื่องานวิจัย บทคัดย่อ คำสำคัญ มาประยุกต์ร่วมกับข้อมูลทางบรรณานุกรมของงานวิจัยแต่ละงานซึ่งประกอบไปด้วย จำนวนการอ้างอิง จำนวนการถูกอ้างอิงถึง และข้อมูลแหล่งการตีพิมพ์งานวิจัย ได้แก่ งานประชุมวิชาการ หรือวารสารวิชาการ จำนวนเอกสารที่มีการตีพิมพ์ในงานประชุมวิชาการหรือวารสารวิชาการ จากผลการทดสอบเบื้องต้น พบว่าการผสมผสานระหว่าง *Query Dependent Ranking* และ *Query Independent Ranking* สามารถให้ผลการค้นคืนเอกสาร ที่มีคุณภาพในห้าลำดับแรกดีกว่าวิธีการสืบค้นแบบใช้ *Query Dependent Ranking* เพียงอย่างเดียว

คำสำคัญ: การเรียงลำดับผลลัพธ์การค้นคืน ตัวแบบ ระบบค้นคืนบทความวิจัย

Abstract

Query Dependent Ranking or *Similarity Ranking* is a technique for ranking search results by comparing query terms with document indexes. This technique doesn't consider other related factors such as quality of documents. While, *Query Independent Ranking* or *Static Ranking* is another important ranking search results technique by focusing on quality of documents. This research paper proposed a model for ranking search results in a research paper search engine. The proposed a technique to combine *Query Dependent Ranking* with *Query Independent Ranking* using bibliographic information. The bibliographic information was used to determine a quality of research papers. This technique started with the usage of similarity feature, such as title of research papers, abstract, and keywords, in combination with a bibliographic information of each research paper, such as number of citation, number of cited by other papers, and source of publication including conference proceeding, peer-review journal. From the preliminary result of experiment, the combination technique between *Query Dependent Ranking* and *Query Independent Ranking* provide more relevant research paper search results for the top five ranking results comparing with the results from the *Query Dependent Ranking* technique only.

Keyword: Ranking Search Results, Model, Research Paper Search Engine

1. บทนำ

เทคโนโลยีสารสนเทศและอินเทอร์เน็ตถูกพัฒนา ไปอย่างรวดเร็ว ทำให้ปริมาณข้อมูลและสารสนเทศต่างๆ ถูกเผยแพร่มากมายมหาศาล ดังนั้นการพัฒนาระบบการสืบค้นข้อมูลที่มีประสิทธิภาพ และตรงกับความต้องการของผู้ใช้จึงทำได้ยากมากขึ้นด้วย

ระบบสืบค้นที่นำมาใช้งานในอดีต เช่น การสืบค้นข้อมูลของ Yahoo! ใช้วิธีการที่เรียกว่า Catalog Based Information Retrieval จะเป็นการสืบค้นจากหมวดหมู่หลัก แล้วย่อยลงไปจนถึงหัวข้อที่ต้องการ หรืออีกวิธีการหนึ่งที่นิยมใช้กันอยู่แพร่หลายก็คือ Query Based Search Engine เป็นการสืบค้นข้อมูลที่มักจะพิจารณา โดยเปรียบเทียบความเหมือนระหว่างคำค้น (Query) และดัชนีของเอกสารเท่านั้น เอกสารที่ได้จากผลลัพธ์การค้นคืนจะไม่มีความสัมพันธ์กับคำค้นที่ต้องการเลย ซึ่งวิธีการดังกล่าวถูกเรียกว่า Query Dependent Ranking หรือ Similarity Ranking

ระบบสืบค้นในยุคถัดมา เช่น กูเกิ้ล (Google) เริ่มมีการนำเอาปัจจัยที่เกี่ยวข้อง อย่างคุณภาพของเอกสารเข้ามาร่วมพิจารณาในการเรียงลำดับผลลัพธ์การค้นคืน โดยวิธีการดังกล่าวถูกเรียกว่า Query Independent Ranking หรือ Static Ranking [1]

ในการศึกษางานวิจัยนี้ทำการทดลองเพื่อพิสูจน์สันนิษฐานว่า เทคนิคสำหรับการเรียงลำดับแบบ Query Dependent Ranking แบบผสมผสานข้อมูลบรรณานุกรม ให้ผลลัพธ์ดีกว่า Query Dependent Ranking เพียงอย่างเดียว โดยการสร้างดัชนีต้นแบบ ของทั้งสองวิธี และเพื่อให้ผลการทดลองสามารถควบคุมปัจจัย ภายในและขอบเขตของข้อมูล ในผลงานชิ้นนี้จึงใช้มุ่งเน้นไปที่ข้อมูลบทความวิจัย

ใน ส่วนที่ 2 กล่าวถึงงานวิจัยที่เกี่ยวข้องกับระบบสืบค้น (Search Engine) และระบบสืบค้นบทความวิจัย (Research Paper Search Engine) ที่ผ่านมา ส่วนที่ 3 การนำเสนอตัวแบบประกอบด้วยภาพรวมของระบบ ตั้งแต่ขั้นตอนการเก็บข้อมูล เพื่อสร้างดัชนีต้นแบบ ใช้ตัวแทนของเอกสารทั้งหมดในระบบสืบค้นที่อยู่ในคลังเอกสาร (Paper Corpus) ส่วนที่ 4 กล่าวถึงวิธีการทดลอง ส่วนที่ 5 การประเมินผลระบบสืบค้นของแต่ละดัชนี และการอภิปรายผลในหัวข้อสุดท้าย

2. งานวิจัยที่เกี่ยวข้อง

ในกระบวนการของระบบสืบค้นสารสนเทศ มีวิธีการนำเสนอที่หลากหลาย ขึ้นอยู่กับลักษณะและประเภทของสารสนเทศ ในช่วงเวลา 5 ปีที่ผ่านมา งานวิจัยที่เกี่ยวข้องมักจะใช้ข้อมูลต่างๆ จากระบบเครือข่ายทางสังคม (Social Networking System) เช่น CiteULike และ del.io.us โดยนำ Social Tag และข้อมูลที่เกี่ยวข้องกับการ Bookmark มาใช้เพื่อทำ Query Dependent Ranking และ Query Independent Ranking [2] นอกจากนี้แล้วยังนำเอาข้อมูลดังกล่าวมาทำ Profile ของผู้ใช้งาน เพื่อสร้างระบบแนะนำเนื้อหา (Recommendation System) [3], [4], [5], [6], [7], [8]

สำหรับงานวิจัยที่เกี่ยวข้องกับการเรียงลำดับผลลัพธ์การค้นคืนของบทความวิจัยนั้น มีการเสนอวิธีการในการเรียงลำดับผลลัพธ์ โดยการใช้เวลา [9] และใช้ข้อมูลอื่น ๆ เช่น จำนวนของกลุ่มผู้ใช้ที่มีการแชร์บทความ จำนวนผู้ที่ชื่นชอบบทความ และเวลา (Time Stamp) ของการแบ่งปันบทความ [10] ซึ่งจะพบว่ายังมีข้อมูลอื่น ๆ ที่น่าสนใจเอามาใช้ในการประเมินคุณภาพของบทความวิจัย เพื่อทำการจัดลำดับผลลัพธ์การค้นคืนอีก

จากวิจัยที่กล่าวมา พบว่าบทความวิจัยที่มีการนำเสนอไว้ก่อนหน้ามีการนำข้อมูลแวดล้อมของเอกสารมาพิจารณาเพิ่ม และถูกนำมาใช้เรียงลำดับผลลัพธ์ โดยมุ่งเน้นไปที่ลักษณะของผู้ใช้งาน (Profile) มากกว่ามุ่งเน้นไปที่คุณภาพของเอกสารที่ได้ ดังนั้นในงานวิจัยนี้จึงเสนอตัวแบบ สำหรับเรียงลำดับผลลัพธ์การค้นคืนในระบบสืบค้นบทความวิจัย ที่มีการผสมผสานระหว่าง Query Dependent Ranking และ Query Independent Ranking โดยจะนำเอาข้อมูลทางบรรณานุกรมมาใช้สำหรับการประเมินคุณภาพของบทความวิจัย โดยจะเน้นที่การประเมินคุณภาพของแหล่งตีพิมพ์ และข้อมูลการอ้างอิงเป็นหลัก

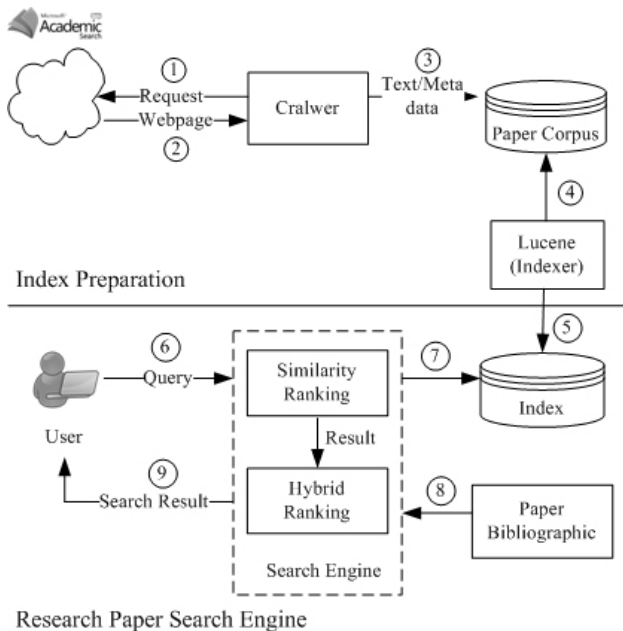
3. การนำเสนอตัวแบบ

การศึกษางานวิจัยมีการสร้างระบบสืบค้นขึ้นเพื่อใช้พิสูจน์ ตัวแบบ มีขั้นตอนการทำงาน แสดงดังภาพที่ 1 ดังนี้

3.1 Crawler

เป็นโปรแกรมที่ทำหน้าที่เก็บข้อมูลจากอินเทอร์เน็ต เพื่อวิเคราะห์ และกรองรายละเอียดที่ต้องเก็บลงฐานข้อมูล โดยใน

การศึกษาใช้ข้อมูลจาก academic.research.microsoft.com ซึ่งเป็นผู้ให้บริการข้อมูลบทความวิจัยสาขาต่างๆ ในแต่ละบทความมีรายละเอียดทางบรรณานุกรม เช่น ชื่อหัวข้อบทความวิจัย (Title) ผู้แต่ง (Author) คำสำคัญ (Keyword) บทคัดย่อ (Abstract) ปีที่ตีพิมพ์ (Year) จำนวนการถูกอ้างอิงถึง (Citation) จำนวนเอกสารอ้างอิง (Reference) และรายละเอียดของงานประชุมวิชาการที่บทความได้ตีพิมพ์ (Publisher) เช่น บทความนี้ถูกตีพิมพ์ในงานประชุมวิชาการ (Conference) หรือวารสารวิชาการ (Journal) จำนวนบทความวิจัยที่ตีพิมพ์แล้ว (Publication) ระยะเวลาปีที่มีการตีพิมพ์ (Year Range) และจำนวนเอกสารมีการอ้างอิงถึง (Publisher Citation)



ภาพที่ 1: ขั้นตอนการทำงานของระบบ

3.2 Paper Corpus

เป็นคลังเอกสารที่เก็บรวบรวมและบันทึกบทความวิจัยทั้งหมดของระบบที่ได้จาก Crawler

3.3 ดัชนี (Index)

ดัชนีเป็นขั้นตอนการเตรียมข้อมูล โดยใช้ Lucene เป็นไลบรารีสำหรับคำนวณค่าความถี่ของคำในเอกสาร (Term) ที่ปรากฏอยู่ในเอกสารทั้งหมด สำหรับการสืบค้นที่มีประสิทธิภาพและรวดเร็ว จัดเก็บอยู่ในรูปแบบของเวกเตอร์ (Vector) วิธีการนี้จะถูกนำมาใช้ในงาน Information Retrieval และ

Text Mining ซึ่งจะเห็นว่าการส่วนใหญ่จะเป็นการเปรียบเทียบเอกสารที่อยู่ใน Corpus กับคำค้นเท่านั้น

คำนวณค่าน้ำหนัก $w_{t,D}$ ของคำ หมายถึงคำแต่ละคำมีความเกี่ยวข้อง (Relevant) กับเอกสารมากน้อยแค่ไหน ได้จาก $tf.idf$ แสดงดังสมการที่ (1) และ (2)

$$tf.idf_{t,D} = tf_{t,D} \times idf_t \quad (1)$$

$$w_{t,D} = tf_{t,D} \times \log\left(\frac{N}{df_t}\right) \quad (2)$$

เมื่อ $w_{t,D}$ แทนค่าน้ำหนักของ $tf.idf_{t,D}$ และหมายถึงความเกี่ยวข้องของคำ t ในเอกสาร D

$tf_{t,D}$ (Term Frequency) แทนจำนวนคำ t ที่ปรากฏในเอกสาร D

N แทนจำนวนเอกสารทั้งหมด

idf_t (Inverse Document Frequency) แทนจำนวนเอกสารที่มี t ปรากฏอยู่ จากสมการ (2) พบว่าเมื่อ t ปรากฏอยู่ในทุกเอกสาร N ส่งผลให้ค่า idf_t มีค่าความสำคัญลดลงจนมีค่าเป็นศูนย์

3.4 Re-Ranking Model

จากสมการที่กล่าวไปก่อนหน้านี้ ทำการทดลองโดยสร้างต้นแบบดัชนีทั้งหมด 4 แบบ ดังนี้

1. Index0 แทน Full-Text Index

2. Index 1 แทน Full-Text Boost Field Index มีการเพิ่มค่าน้ำหนักให้กับฟิลด์ข้อมูล (Boost Field) ได้แก่ ชื่อ (Title) บทคัดย่อ (Abstract) และ คำสำคัญ (Keyword) ค่าน้ำหนักแต่ละตัวเป็น 3 2 และ 1 ตามลำดับ

3. Hybrid0 แทน Full-Text Index กับ Bibliographic

4. Hybrid1 แทน Full-Text Boost Field Index ได้แก่ ชื่อ (Title) บทคัดย่อ (Abstract) และ คำสำคัญ (Keyword) ค่าน้ำหนักแต่ละตัวเป็น 3 2 และ 1 ตามลำดับ กับ Bibliographic

เมื่อ Index0 และ Index1 ใช้คุณสมบัติจาก Similarity Feature และ Hybrid0 และ Hybrid1 ใช้ Similarity Feature ร่วมกับ Bibliographic Feature สามารถหาค่าน้ำหนักของเอกสารดังนี้

ความสัมพันธ์ระหว่าง Similarity Feature กับ Bibliographic Feature ของเอกสารงานวิจัยและผู้จัดพิมพ์ ตามสมการที่ (3)

$$Hybrid\ Score = Sim(\alpha) + Bib(1 - \alpha) \quad (3)$$

เมื่อ Hybrid Score แทนค่าคะแนนที่เกิดจากการวัดคุณภาพของ Similarity Score ร่วมกับ Bibliographic Score Sim แทน Similarity Score

Bib แทน Bibliographic Score คัดจากค่าเฉลี่ยจากการวัดค่าคุณภาพของบทความวิจัย ร่วมกับคุณภาพของผู้จัดพิมพ์งานวิจัยได้จากสมการที่ (4) และ (5) ตามลำดับ

$$QA = R(\beta) + CA(1 - \beta) \quad (4)$$

QA แทนคุณภาพของเอกสารงานวิจัย (Article Quality) ประกอบด้วย R และ CA และแทนค่า β เท่ากับ 0.9 ค่าน้ำหนักที่กำหนดให้

R แทนจำนวนเอกสารอ้างอิงภายในบทความวิจัยนั้น ค่าที่นำมาใช้ทำ Scale Normalized อยู่ระหว่าง 0 ถึง 1 คำนวณจากเอกสาร 30 เอกสารแรกที่ได้จากการค้นคืน

CA แทนจำนวนเอกสารที่มีการอ้างอิงถึงบทความวิจัยนั้น ค่าที่นำมาใช้ทำ Scale Normalized อยู่ระหว่าง 0 ถึง 1 คำนวณจากเอกสาร 30 เอกสารแรกที่ได้จากการค้นคืน

$$QP = Type \times CP \quad (5)$$

QP แทนคุณภาพของผู้จัดพิมพ์งานวิจัย (Publisher Quality)

Type แทนประเภทของผู้จัดพิมพ์ กำหนดให้วารสารวิชาการ (Journal) เท่ากับ 1.0 งานประชุมวิชาการ (Conference) เท่ากับ 0.1

CP แทนจำนวนเอกสารที่มีการอ้างอิงถึงผู้จัดพิมพ์ เปรียบเทียบกับจำนวนเอกสารที่ตีพิมพ์ของผู้จัดพิมพ์นี้ และค่าที่นำมาใช้ต้อง Scale Normalized อยู่ระหว่าง 0 ถึง 1

4. วิธีการดำเนินการวิจัย

4.1 การเตรียมคลังเอกสารงานวิจัย

เอกสารงานวิจัยจาก academic.research.microsoft.com รวบรวมระหว่างเดือนมิถุนายนถึงสิงหาคม ปี 2013 ประกอบด้วยงานวิจัยจำนวน 71,828 บทความ โดยจำแนกเป็นบทความจากวารสารวิชาการ (Journal) 28,320 บทความ และบทความจากงานประชุมวิชาการ (Conference) 43,508 บทความ และในคลังเอกสาร (Document Corpus) ประกอบด้วยมีคำสำคัญ (Keyword) 23,073

งานวิจัยประกอบด้วย ชื่อ รายชื่อนักวิจัย บทคัดย่อ คำสำคัญ (Keyword) จำนวนเอกสารอ้างอิง (Reference) จำนวนผลงานถูกอ้างอิง (Citation) ปีและผู้จัดพิมพ์งานวิชาการ (Publisher) เช่น วารสารวิชาการ (Journal) 4,283 หรืองานประชุมวิชาการ (Conference) 2,885

4.2 การประเมินผล

วิธีการวัดประสิทธิภาพของ Indexing และ Ranking คือ ค่า Normalized Discounted Cumulative Gain (NDCG) โดย Jarvelin, Kekalainen [9]

การประเมินผลมาจากคะแนนผู้ใช้เป็นหลัก (Judgments) เรียกว่า Perfect Score นำมาคำนวณเป็น DCG Perfect คะแนนที่ได้ บ่งบอกว่าคำค้นมีความเกี่ยวข้องกับเอกสารนั้นๆ ที่ตำแหน่งที่ k เมื่อกำหนดให้คำค้น q และเซตเอกสารจากการค้นคืน คะแนนของเอกสารในแต่ละตำแหน่งสามารถคิดได้จากลำดับแรกจนถึงเอกสารลำดับสุดท้าย ตามสมการที่ (6)

$$NDCG_q = \frac{\sum_{j=1}^k (2^{r(j)} - 1)}{\log(1 + j)} \quad (6)$$

เมื่อ j แทนตำแหน่งของเอกสาร และ $r(j)$ แทนเลขจำนวนเต็ม ซึ่งเป็นค่าคะแนน (Judgment Score) ที่ได้จากผู้ทดสอบ

NDCG แทนค่าคะแนนความเกี่ยวข้องของเอกสารจากลำดับแรกสุด ไปยังลำดับสุดท้าย

4.3 การทดสอบ

การทดสอบเพื่อพิสูจน์ตัวแบบที่นำเสนอ จึงจัดทำระบบ Research Paper Search Engine สำหรับเป็นหน้าเว็บของระบบสืบค้นที่เป็น Interface ให้กับผู้ทดสอบ

ในการทดสอบได้เชิญนักศึกษาระดับปริญญาโท ปริญญาเอก และนักวิจัยด้านวิทยาการคอมพิวเตอร์ กำหนดให้ผู้ทดสอบแต่ละคนใส่คำค้นที่ต้องการ เป็นคำ หรือประโยคใดๆ ก็ได้ในหน้าเว็บ ระบบจะสืบค้นข้อมูลจากทั้ง 4 ดัชนี โดยก่อนที่จะแสดงผลให้ผู้ทดสอบประเมิน ระบบจะมีการสุ่มลำดับและรวมผลลัพธ์ที่ได้เพื่อไม่ให้มีการแสดงเอกสารซ้ำในแต่ละดัชนี และเพื่อให้ผู้ทดสอบเกิดความลำเอียงในการให้คะแนนเอกสารที่ได้ในแต่ละลำดับด้วย

หน้าเว็บที่แสดงผลลัพธ์ โดยแสดงรายละเอียดบทความวิจัยได้แก่ ชื่อ และบทคัดย่อ โดยผู้ทดสอบจะต้องอ่านรายละเอียดทั้งในส่วนของหัวข้อและบทคัดย่อ แล้วให้คะแนนบทความที่กำลังพิจารณาว่ามีความเกี่ยวข้องกับคำค้นมากน้อยแค่ไหน ซึ่งคะแนนที่ได้ดังกล่าวจะนำไปประเมินผลการเรียงลำดับที่ได้ (Research Paper Re-Ranking)

การทดสอบมีขั้นตอน ดังนี้

1. ผู้ทดสอบระบุคำค้นที่ต้องการในหน้าเว็บ
2. ระบบจะสืบค้นเอกสาร 30 ลำดับแรก ของแต่ละดัชนี โดยระบบจะตรวจสอบเอกสารที่แสดงผลซ้ำ และสุ่มลำดับการแสดงผลบนหน้าเว็บ เพื่อไม่ให้ผู้ทดสอบเกิดความลำเอียงในการให้คะแนนในแต่ละเอกสาร
3. ผู้ทดสอบให้คะแนนเอกสาร (Judgment Score) แต่ละเอกสารมีความเกี่ยวข้องกับคำค้นอย่างไร คะแนนอยู่ระหว่าง 4 ถึง 0 มีความหมาย ดังตารางที่ 1
4. ระบบบันทึกข้อมูล

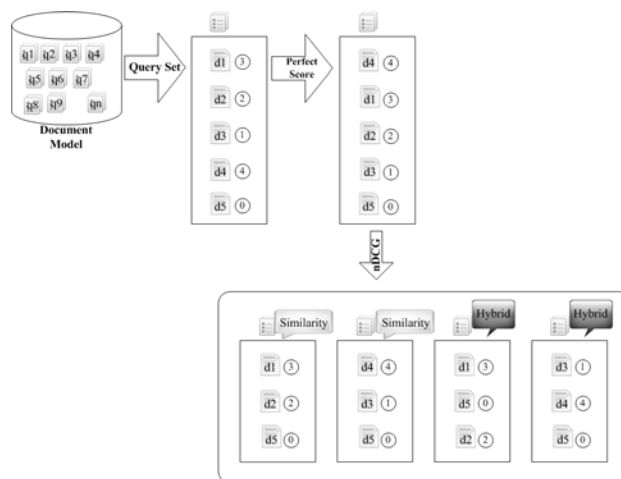
ตารางที่ 1 : Judgment Score

คะแนน	รายละเอียด
4	มีความเกี่ยวข้อง (Extremely Relevant)
3	อาจจะเกี่ยวข้อง (Probably Relevant)
2	เกี่ยวข้องเพียงเล็กน้อย (Less Relevant)
1	อาจจะไม่เกี่ยวข้อง (Probably Not Relevant)
0	ไม่มีความเกี่ยวข้อง (Not Relevant)

5. ผลการทดลอง

จาก Judgment Score ที่ได้จากผู้ทดสอบ ในแต่ละครั้งของการสืบค้น คือข้อมูล 1 ชุดประเมิน แสดงดังภาพที่ 2 เอกสาร

ทั้งหมดในหนึ่งชุดจะถูกเรียงลำดับตาม Judgment Score เพื่อหาค่า DCG Perfect Score และ ขั้นตอนที่ 2 เป็นการจำแนกเอกสารออกเป็น Index0 Index1 Hybrid0 และ Hybrid1 ได้ 4 ชุดดัชนี สำหรับหาค่า NDCG ที่ตำแหน่งการค้นคืนเอกสารที่ k



ภาพที่ 2: การคำนวณค่า NDCG

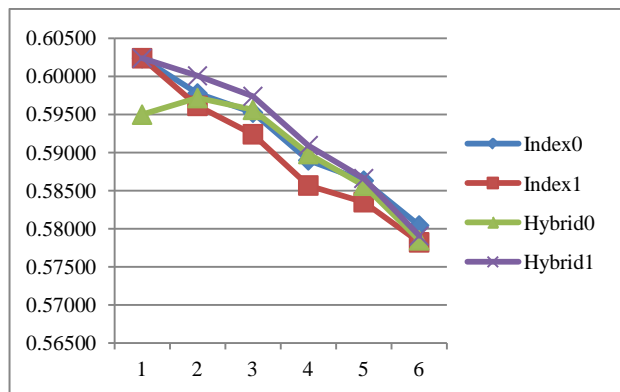
จากผลการทดสอบเบื้องต้น มีผู้ร่วมทดสอบ 20 คน โดยมีจำนวนคำค้นทั้งหมด 37 คำสืบค้น ในการประเมินผลหาค่าเฉลี่ย NDCG ในแต่ละดัชนี เมื่อแกน x แทนลำดับผลการค้นคืนใน 6 ลำดับแรก และแกน y แทนค่าเฉลี่ยของ NDCG ที่ได้จากการสืบค้น

เมื่อพิจารณาจากกราฟ แสดงดังภาพที่ 3 พบว่า NDCG ที่ตำแหน่งเอกสาร $k=1$ ที่ Index0 และ Hybrid1 ได้ค่า NDCG เท่ากับ 0.60238 เท่ากัน ดังตารางที่ 2 แต่เมื่อพิจารณาค่าเฉลี่ย NDCG ในช่วงตำแหน่งที่ $k=2$ ถึง 4 ค่าเฉลี่ย NDCG ของดัชนี Hybrid1 สามารถให้ผลการค้นคืนเอกสารที่ดีที่สุด โดยพิจารณาจากค่าเฉลี่ย NDCG ของผลลัพธ์

ค่าดัชนีของ Hybrid1 มีการกำหนด α เท่ากับ 0.9 ตามสมการที่ (3) ซึ่งเป็นอัตราส่วนระหว่างค่าน้ำหนักของ Similarity กับ Static Ranking ที่มีส่วนช่วยให้อัตราการเรียงลำดับผลลัพธ์การค้นคืนดีขึ้น เมื่อเปรียบเทียบกับการเรียงลำดับบทความวิจัย ซึ่งส่วนใหญ่จะให้ความสำคัญกับความเหมือนมากกว่าคุณภาพของเอกสาร

ค่าน้ำหนักที่ได้จากการทดลอง เป็นผลของการนำเอา Static Feature มาช่วยในการเรียงลำดับ ส่งผลให้เกิดประสิทธิภาพของ

การเรียงลำดับที่ดีขึ้นเมื่อนำมาใช้ร่วมกับ Similarity Feature เพียงอย่างเดียว



ภาพที่ 3: เปรียบเทียบค่า NDCG ของแต่ละดัชนี

ตารางที่ 2 ค่าเฉลี่ย NDCG

k	Index0	Index1	Hybrid0	Hybrid1
1	0.60238	0.60238	0.59497	0.60238
2	0.59776	0.59613	0.59715	0.60007
3	0.59524	0.59240	0.59561	0.59739
4	0.58898	0.58567	0.58980	0.59093
5	0.58630	0.58349	0.58565	0.58660
6	0.58042	0.57821	0.57848	0.57905

6. สรุปและอภิปรายผล

จากงานทดสอบเบื้องต้นพบว่า เอกสารงานวิจัยที่มีจำนวนการอ้างอิงจากเอกสารอื่นมากกว่า เป็นงานที่มีการตีพิมพ์หรือเผยแพร่ออกมาแล้วในช่วงระยะเวลาหนึ่ง ซึ่งเป็นจุดที่ทำให้งานวิจัยใหม่หรืองานที่เพิ่งตีพิมพ์ออกมาปัจจุบันยังไม่สามารถถูกค้นขึ้นมาแสดงผลการค้นคืนในลำดับต้นๆ ได้ ดังนั้นกรอบแนวความคิดนี้ จึงเหมาะสำหรับการสืบค้นเอกสารที่เน้นด้านคุณภาพมากกว่าความใหม่ของงานวิจัย เช่น ปีที่ตีพิมพ์ ซึ่งเป็นปัจจัยที่แปรผกผันกับจำนวนการถูกอ้างอิงจากเอกสารอื่นๆ

ค่า Judgment Score ที่ใช้เป็นพื้นฐานสำหรับคำนวณค่าที่บ่งบอกถึงความพึงพอใจของผู้ใช้งาน จำเป็นต้องเพิ่มปริมาณผู้ทดสอบเพื่อให้ได้ค่าเฉลี่ยที่ถูกต้องและผิดพลาดน้อยลง และสามารถแสดงผลค่าทางสถิติได้อย่างมีนัยสำคัญต่อไป

เอกสารอ้างอิง

- [1] X.Geng, T. Y. Liu, and H. Li, "Feature selection for ranking," *Proc ACM SIGIR*, 2007.
- [2] W. Choochaiwattana, M. B. Spring, "Applying social annotation to retrieve and re-rank web resources," *International conference on information management and engineering (IEEE)*, 2009.
- [3] W.Choochaiwattana, "Usage of tagging for research paper recommendation," *3rd International conference on Advanced computer theory and engineering (IEEE)*, 2010.
- [4] S.Noel, R.Beale, "Sharing vocabularies: tag usage in CiteULike," *the British computer society*, 2008.
- [5] D.H.Lee, P.Brusilovsky, "Socail networks and interest similarity: the case of CiteULike," *ACM*, 2010.
- [6] H.Wang, X.He, M.Chang, Y.Song, R.W.White, W.Chu, "Personalized ranking model adaptation for web search" *ACM*, 2013
- [7] D.Parra, P.Brusilovsky, "Collaborative filtering for social tagging system: an experiment with CiteULike," *ACM*, 2009.
- [8] M.S.Pera, Y.Ng, "A personalized recommendation system on scholarly publications," *ACM*, 2011.
- [9] P. Jomsri, "A combination of similarity ranking and time for social research paper searching," *World academy of science, engineering and technology 54*, 2011.
- [10] P. Jomsri, S. Sanguansintukul, and W. Choochaiwattana, "CiteRank: combination similarity and static ranking with research paper searching," *International Journal of Internet Technology and Secured Transactions*, Volume 3 Issue 2, April 2011 .
- [11] K. Jarvelin, and J. Kekalainen, "IR evaluation methods for retrieving highly relevant documents," *Proc. ACM SIGIR conference on Research and Development on Information Retrieval*, July 2000.
- [12] Q. Wu, C. J. C. Burges, K. M. Svore and J Gao, "Microsoft research technical report MSR-TR-2008-109," October 15, 2008.

กลไกความปลอดภัยของเอกสาร WSDL ในเว็บเซอร์วิส

A Mechanism for Securing WSDL Web Service

เจนจิรา หวังหลี (Janejira Wanglee)¹ และ ลัดดา ปรีชาวีรกุล (Ladda Preechaveerakul)²
 ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตหาดใหญ่
 ตำบลคอหงส์ อำเภอหาดใหญ่ สงขลา 90110
¹janeji.184@gmail.com, ²ladda.p@psu.ac.th

บทคัดย่อ

ปัจจุบันมีการใช้งานเว็บเซอร์วิสเพิ่มมากขึ้นบนอินเทอร์เน็ต และการเรียกใช้งานเว็บเซอร์วิสนั้นจะมีเอกสาร WSDL (Web Service Description Language) ซึ่งเป็นเอกสารที่อธิบายวิธีการเรียกใช้งานเว็บเซอร์วิส ทำให้มีโอกาสเกิดการโจมตีเอกสาร WSDL ขึ้นเพื่อขโมยข้อมูล โดยทั่วไปเพื่อให้เอกสารมีความปลอดภัยจะเข้ารหัสเอกสาร WSDL ด้วยวิธีการมาตรฐานของ XML อย่างไรก็ตามการเข้ารหัสและถอดรหัสเอกสารด้วยวิธีนี้ใช้ระยะเวลาพอสมควร งานวิจัยนี้จึงนำเสนอกลไกความปลอดภัยของเอกสาร WSDL ในเว็บเซอร์วิส โดยการเข้ารหัสเอกสาร WSDL เพียงบางส่วน ซึ่งผลการทดลองแสดงให้เห็นว่า กลไกที่นำเสนอลดระยะเวลาในการเข้ารหัสได้ถึง 37.74% และลดระยะเวลาในการถอดรหัสได้ถึง 22.50% เมื่อเทียบกับงานวิจัยก่อนหน้า

คำสำคัญ: XML WSDL ความปลอดภัย เว็บเซอร์วิส

Abstract

At present, web services are widely used over the Internet. Since Web Service Description Language (WSDL) is used to describe the public interface to a specific web service, it is possible to be attacked from an unauthorized user. Generally, a WSDL document is encrypted using XML security standards. However, it spends times to encrypt and decrypt the WSDL document. This paper then, proposes a mechanism for securing WSDL documents in web service (SWSDL) using partial encryption in the WSDL file. The experimental result shows that SWSDL can reduce the encryption time to 37.74% and the decryption time to 22.50% in comparison with a previous research.

Keywords: XML, WSDL, Security, Web service

1. บทนำ

ปัจจุบันนี้มีการนำเว็บเซอร์วิสมาประยุกต์ใช้งานเพื่อตอบสนองความต้องการทางด้านธุรกิจ ซึ่งเว็บเซอร์วิสสนับสนุนการทำงานระหว่างกัน (Interoperation) โดยใช้ภาษา XML [1] ในการรับส่งข้อมูลระหว่างผู้ให้บริการและผู้ใช้บริการ ดังนั้นหลายองค์กรจึงหันมาให้ความสำคัญกับการใช้งานเว็บเซอร์วิสมากขึ้น อย่างไรก็ตามการใช้งานเว็บเซอร์วิส ผู้ให้บริการเซอร์วิส นั้นๆจำเป็นต้องรู้วิธีการติดต่อกับผู้ให้บริการ ซึ่งวิธีการติดต่อเรียกใช้งานเซอร์วิสอธิบายในเอกสาร WSDL (Web Service Description Language) [2] การใช้งานเว็บเซอร์วิสโดยทั่วไปไม่มีกลไกสำหรับความปลอดภัยเพื่อเผยแพร่ข้อมูลเอกสาร WSDL ผ่านเครือข่ายอินเทอร์เน็ตจากผู้ให้บริการไปยังผู้ให้บริการ ซึ่งเอกสาร WSDL อาจถูกโจมตีได้ แม้ว่า XML จะมี XML Encryption [3] และ XML Signature [3] เป็นมาตรฐานรองรับความปลอดภัยในส่วนนี้แล้วก็ตาม แต่ด้วยขั้นตอนวิธีที่ยุ่งยากซับซ้อนและใช้เวลาพอสมควรในการเข้ารหัสและถอดรหัสข้อมูลเพียงบางส่วนเพื่อต้องการปกปิดข้อมูลในเอกสาร WSDL ให้เป็นความลับ บทความนี้จึงเสนอกลไกสำหรับความปลอดภัยของเอกสาร WSDL โดยการเข้ารหัสเอกสาร WSDL เพียงบางส่วน เพื่อปกปิดข้อมูล และป้องกันการปลอมแปลงเอกสาร WSDL โดยผู้ให้บริการถอดรหัสเอกสาร WSDL ก่อนเรียกใช้เว็บเซอร์วิสนั้น กลไกดังกล่าวทำให้ข้อมูลเอกสาร WSDL เป็นความลับและมีความปลอดภัยเมื่อส่งผ่านระบบเครือข่ายอินเทอร์เน็ตจากผู้ให้บริการไปยังผู้ให้บริการ ป้องกันการโจมตีเอกสาร WSDL และลดระยะเวลาในการเข้ารหัสและถอดรหัสข้อมูล ทำให้การใช้งานเว็บเซอร์วิสมีความปลอดภัยและน่าเชื่อถือ

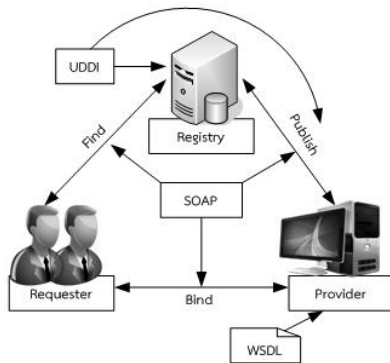
เนื้อหาของบทความในส่วนที่ 2 กล่าวถึงทฤษฎีและงานวิจัยที่เกี่ยวข้อง ส่วนที่ 3 อธิบายการทำงานของกลไกความปลอดภัย

ของเอกสาร WSDL ในส่วนที่ 4 แสดงการพัฒนาทั่วโลกความปลอดภัยและผลลัพธ์ และส่วนที่ 5 เป็นบทสรุป

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 เว็บเซอร์วิส

เว็บเซอร์วิส [4] ประกอบด้วย WSDL SOAP (Simple Object Access Protocol) และ UDDI (Universal Description Discovery and Integration) แสดงดังภาพที่ 1 โดยมีการทำงานร่วมกันดังนี้



ภาพที่ 1: พื้นฐานเว็บเซอร์วิส [2]

ผู้ให้บริการ .1(Provider) เตรียมบริการเซอร์วิสให้เรียกใช้และเผยแพร่เซอร์วิสที่จัดทำขึ้นซึ่งก็คือ เอกสาร WSDL ไปยัง Registry หรือ UDDI

2. UDDI ให้บริการลงทะเบียน (Registry) เซอร์วิสและค้นหาเซอร์วิส UDDI เป็นบริการเก็บรวบรวมเซอร์วิสที่ผู้ให้บริการประกาศไว้

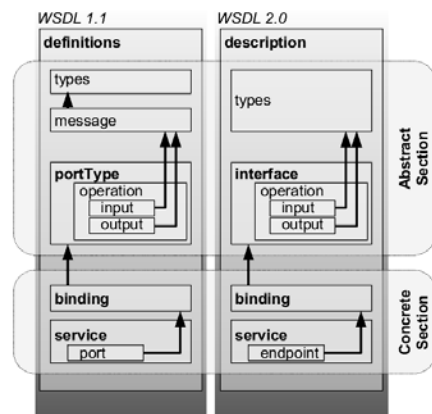
3. ผู้ขอใช้บริการ (Requester) เป็นผู้เรียกใช้เซอร์วิส ค้นหาเซอร์วิสจาก Registry แล้วเรียกใช้เซอร์วิสจากผู้ให้บริการ ซึ่งใช้ SOAP เป็นโพรโทคอลในการแลกเปลี่ยนข้อมูลระหว่างกันบนเว็บเซอร์วิส

องค์กร W3C [5] เป็นผู้กำหนดมาตรฐานเอกสาร WSDL ซึ่งเป็นเอกสารที่อยู่ในรูปแบบของภาษา XML เพื่ออธิบายการเรียกใช้งานเว็บเซอร์วิส ตัวอย่างและคำอธิบายแต่ละแท็กของเอกสาร WSDL 1.1 แสดงดังตารางที่ 1

ตารางที่ 1: ตัวอย่างแท็กของเอกสาร WSDL1.1

WSDL tag	คำอธิบาย
<definitions>	กำหนดชื่อของเซอร์วิส
<types>	อธิบายชนิดข้อมูล
<message>	อธิบายข้อมูลที่ใช้ในการแลกเปลี่ยน
<portType>	กำหนดฟังก์ชันเรียกใช้งาน
<binding>	อธิบายโพรโทคอล
<service>	กำหนดที่อยู่ของเซอร์วิส

ปัจจุบัน W3C ได้ระบุข้อกำหนดเกี่ยวกับ WSDL2.0 ซึ่งมีข้อแตกต่างจาก WSDL1.1 ดังภาพที่ 2



ภาพที่ 2: เปรียบเทียบเอกสาร WSDL 1.1 กับ WSDL 2.0 [6]

2.2 การโจมตีเอกสาร WSDL

ปัจจุบันนี้มีการโจมตีเอกสาร WSDL [7, 8] แบ่งเป็น 2 ประเภทหลักๆคือ 1) *WSDL Scanning* เป็นการโจมตีโดยผู้โจมตี (Hacker) อ่านแล้ววิเคราะห์เอกสาร WSDL และหาช่องโหว่ในเอกสาร WSDL เพื่อดึงข้อมูลจากเว็บเซอร์วิส 2) *Parameter Tampering* เป็นการใช้งานเอกสาร WSDL ผิดรูปแบบ โดยผู้โจมตีส่งพารามิเตอร์ซ้ำๆกันหลายครั้ง แล้วหารูปแบบพารามิเตอร์ที่แตกต่างกันในการเข้าถึงข้อมูลที่ไม่ได้รับอนุญาต เช่น การใส่อักขระพิเศษเครื่องหมาย ‘ หรือ % เป็นต้น

ซึ่งการโจมตีทั้ง 2 ประเภทนี้โดยส่วนใหญ่จะโจมตีในแท็ก <message> แท็ก <portType> และแท็ก <service> ในเอกสาร WSDL

2.3 งานวิจัยที่เกี่ยวข้อง

เว็บเซอร์วิสถูกนำมาใช้งานอย่างแพร่หลายในด้านธุรกิจ การทำธุรกรรมทางการเงิน เมื่อมีผู้ประสงค์ร้ายต้องการ โจรกรรมข้อมูล จึงเกิดการโจมตีเว็บเซอร์วิสในส่วนของ เอกสาร WSDL และได้มีงานวิจัยที่พัฒนาวิธีการป้องกันการ โจมตีเอกสาร WSDL ในเว็บเซอร์วิส

ปี ค.ศ. 2007 Sidharth และ Liu [9] ได้เสนอวิธีการจำกัดผู้ใช้ ในการเข้าถึงเอกสาร WSDL และเอกสาร WSDL ไม่มีข้อมูลที่ เป็นประโยชน์ต่อผู้โจมตี

ปี ค.ศ. 2011 Mirtalebi และ Khayyambashi [10] ได้เสนอ วิธีการเข้ารหัสแท็ก <message> และแท็ก <portType> ด้วย XML Encryption โดยใช้กุญแจสมมาตรหรือกุญแจอสมมาตร ขึ้นอยู่กับความเหมาะสมในการนำเว็บเซอร์วิสไปใช้งานและปี ค.ศ 2012 [11] ได้พัฒนาโดยนำ XML Signature มาประยุกต์ เพื่อให้ WSDL ปลอดภัยมากขึ้น

Shahgholi และคณะ [12] ได้เสนอการเข้ารหัสเอกสาร WSDL ด้วย XML Signature และ XML Encryption และ PKI (Public Key Infrastructure) [3] ในปี ค.ศ. 2011

อย่างไรก็ตามงานวิจัยดังกล่าวข้างต้น มีวิธีการเข้ารหัสที่ ยุ่งยากซับซ้อนและใช้ระยะเวลาในการเข้ารหัสพอสมควร ดังนั้นงานวิจัยนี้จึงเสนอกลไกความปลอดภัยของเอกสาร WSDL โดยการเข้ารหัสเอกสาร WSDL เพียงบางส่วนเพื่อลด ระยะเวลาการทำงาน และป้องกันการโจมตีเอกสาร WSDL

3. กลไกความปลอดภัยของเอกสาร WSDL

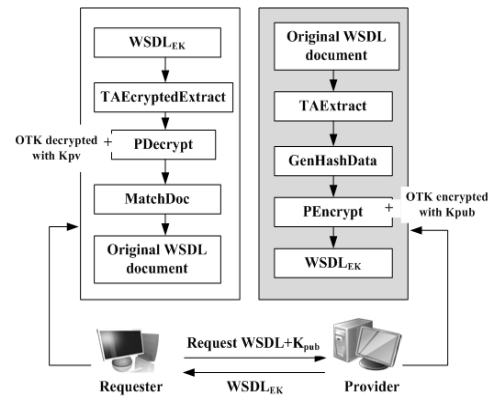
3.1 แนวคิดสำหรับการออกแบบ

เพื่อให้เอกสาร WSDL มีความปลอดภัยสำหรับการนำไปใช้ งาน งานวิจัยนี้จึงเสนอกลไกการเข้ารหัสและถอดรหัสเอกสาร เพียงบางส่วนเพื่อลดระยะเวลา นอกจากนี้ผู้ให้บริการยังได้รับ เอกสาร WSDL ที่เป็นความลับและถูกต้องสมบูรณ์จากผู้ ให้บริการ

3.2 กลไกความปลอดภัยของเอกสาร WSDL (SWSDL)

การทำงานของกลไกความปลอดภัยของเอกสาร WSDL ใน เว็บเซอร์วิส ประกอบด้วยส่วนสำคัญ 2 ส่วน คือ ส่วนเข้ารหัส

เอกสาร WSDL สำหรับผู้ให้บริการ และส่วนถอดรหัสเอกสาร WSDL สำหรับผู้ใช้บริการ ซึ่งกลไกความปลอดภัยของเอกสาร WSDL อธิบายได้ดังภาพที่ 3



ภาพที่ 3: ขั้นตอนของกลไกความปลอดภัยของเอกสาร WSDL(SWSDL)

จากภาพที่ สามารถอธิบายขั้นตอนการทำงานของแต่ละ 3 ส่วนได้ดังนี้

3.2.1. ส่วนเข้ารหัสเอกสาร สำหรับผู้ให้บริการ

เมื่อผู้ใช้บริการ (Requester) ร้องขอเอกสาร WSDL และส่ง กุญแจสาธารณะของตนเองกับผู้ให้บริการ (Provider) ดังนั้นจะ เข้าสู่ส่วนเข้ารหัสเอกสาร WSDL โดยมีขั้นตอนคือ 1) สกัดค่า แอททริบิวต์ในเอกสาร WSDL (TAEExtract) โดยจะสกัดเพียง บางแท็ก คือ แท็ก <message> แท็ก <portType> และ แท็ก <service> 2) กำหนดหมายเลขแต่ละค่าแอททริบิวต์ที่สกัดได้ และนำมาผ่านกระบวนการแฮช (GenHashData) เพื่อย่อ ข้อมความ 3) เข้ารหัสค่าแอททริบิวต์ที่ผ่านกระบวนการกำหนด หมายเลขแล้วโดยใช้กุญแจสมมาตรที่ใช้เพียงครั้งเดียว (One-Time Key: OTK) และเข้ารหัสกุญแจสมมาตรด้วยกุญแจ สาธารณะ K_{pub} (ของผู้ให้บริการ (PEncrypt))

จากนั้นผู้ให้บริการส่งเอกสาร WSDL ที่เข้ารหัสแล้ว (WSDL_{EK}) ซึ่งแนบกุญแจที่ใช้ในการเข้ารหัสไว้ในเอกสาร WSDL ไปยังผู้ใช้บริการ

3.2.2. ส่วนถอดรหัสเอกสาร สำหรับผู้ใช้บริการ

ผู้ใช้บริการถอดรหัสเอกสาร WSDL_{EK} โดยมีขั้นตอนการ ทำงานคือ 1) สกัดแท็กในส่วนของการเข้ารหัสเอกสาร WSDL_{EK} (TAEryptedExtract) 2) ถอดรหัส OTK ด้วยกุญแจ

ส่วนตัว (K_{pv}) ของผู้ให้บริการ และใช้กุญแจนี้ในการถอดรหัสเอกสาร $WSDL_{EK}$ (PDecrypt) 3) นำค่าแอททริบิวต์ที่ได้จากการถอดรหัสผ่านกระบวนการแฮชมาเปรียบเทียบกับค่าแอททริบิวต์ที่สกัดได้จากเอกสาร $WSDL_{EK}$ (MatchDoc) ถ้าผลลัพธ์ที่ได้เท่ากัน ผู้ให้บริการจะนำเอกสาร WSDL ต้นฉบับไปเรียกใช้งานเว็บเซอร์วิสต่อไป

4. การพัฒนากลไกความปลอดภัยและผลลัพธ์

สำหรับกลไกความปลอดภัยของเอกสาร WSDL พัฒนาด้วยภาษาจาวา และประกอบด้วยการทำงานที่สำคัญ ส่วน คือ ส่วน 2 การเข้ารหัสและส่วนการถอดรหัส ภาพที่ 4 แสดงขั้นตอนวิธีการเข้ารหัส

Algorithm1: Encryption data (WSDL)

1. Extract each attribute value (Attr) in <message>, <portType> and <service> tag
2. **For** each Attr
3. Assign number in each Attr ($Attr_{num}$)
4. $Hash_Value = H(Attr_{num})$
5. Replace Attr with Hash_Value
6. Concatenate $Attr_{num}$ with “|” symbol (Msg)
7. **Endfor**
8. Encrypt Msg with OTK
9. Encrypt OTK with K_{pub}
10. Create <wsl:cdescription> tag
11. Insert cipher msg in <wsl:cipher> tag
12. Insert cipher key in <wsl:key> tag

ภาพที่ 4: ขั้นตอนวิธีการเข้ารหัส

จากขั้นตอนวิธีการเข้ารหัสดังภาพที่ 4 ประกอบด้วยขั้นตอนวิธีย่อย 2 วิธีคือ

1. ขั้นตอนการกำหนดหมายเลขและแฮช (GenHashDoc)

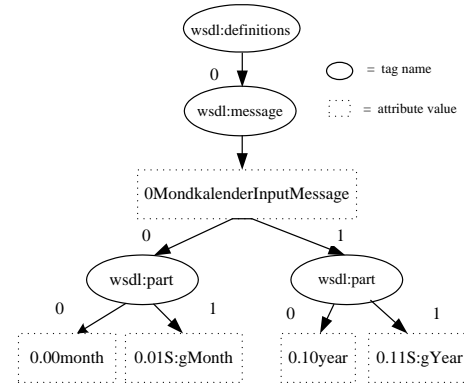
เป็นการกำหนดหมายเลขให้แต่ละค่าแอททริบิวต์ โดยเอกสาร WSDL จะถูกแปลงให้อยู่ในรูปของโครงสร้างต้นไม้ แล้วจึงกำหนดหมายเลขค่าแอททริบิวต์จากบนลงล่างและจากซ้ายไปขวาตามลำดับ ทำให้แต่ละค่าแอททริบิวต์มีหมายเลขไม่ซ้ำกัน เพื่อป้องกันค่าแอททริบิวต์ที่เหมือนกันเมื่อผ่านกระบวนการแฮช

```

<wsl:definitions>
<wsl:message name="MondkalenderInputMessage">
  <wsl:part name="month" type="s:gMonth"/>
  <wsl:part name="year" type="s:gYear"/>
</wsl:message>
</wsl:definitions>
    
```

ภาพที่ 5 : ตัวอย่างแท็ก <wsl:message> ของเอกสาร WSDL1.1

จากตัวอย่างแท็กในภาพที่ 5 เอกสาร WSDL จะนำมาแปลงเป็นรูปแบบโครงสร้างต้นไม้และกำหนดหมายเลขค่าแอททริบิวต์ ($Attr_Id$ และ $Attr_Value$) ดังภาพที่ 6 นำแต่ละแอททริบิวต์ที่สร้างขึ้นมามีผ่านกระบวนการแฮช ได้ $Hash_Value$ และนำไปแทนที่ค่าแอททริบิวต์เดิมในเอกสาร WSDL ดังภาพ 7



ภาพที่ 6 : รูปแบบโครงสร้างต้นไม้ของแท็ก <wsl:message>

```

<wsl:message name="CBAE4DA3FF4C70B3032A6B68879F00661CECE4DD">
  <wsl:part name="F74F4A7582C301F4B2394398DBF21E17B4DC1525" type="8A31279A6C28921CB96784873FDF2BA3C28C242E"/>
  <wsl:part name="DA92C4644F8E821948A290D7E7FD1942B523EC7D" type="119706A849E07067AC74E2E7FDE0CB73DF05EC3A"/> </wsl:message>
    
```

ภาพที่ 7: เอกสาร WSDL ที่ผ่านกระบวนการแฮช

2. ขั้นตอนการเข้ารหัสเอกสาร WSDL (PEncrypt)

เป็นการนำข้อความที่ได้จากการรวมแต่ละค่าแอททริบิวต์ที่กำหนดหมายเลขแล้วมาเข้ารหัสโดยใช้ OTK ด้วยขั้นตอนวิธี AES และเข้ารหัส OTK โดยใช้กุญแจสาธารณะของผู้ให้บริการด้วยขั้นตอนวิธี RSA โดยกำหนดให้เอกสาร $WSDL_{EK}$ ระบุส่วน

ที่เข้ารหัสในเอกสาร WSDL ในแท็ก <wsdl:cdescription> โดยมีรูปแบบดังแสดงในภาพที่ 8 และตัวอย่างแสดงดังภาพที่ 9

```
<wsdl:cdescription>
<wsdl:algorithm> Encryption Algorithm</wsdl:algorithm>
<wsdl:key>Cipher Key</wsdl:key>
<wsdl:cipher>Cipher Message </wsdl:cipher>
</wsdl:cdescription>
```

ภาพที่ 8: รูปแบบแท็ก <wsdl:cdescription> ของเอกสาร WSDL

จากภาพที่ 8 กำหนดให้

<wsdl:cdescription> คือ ข้อมูลที่จำเป็นเกี่ยวกับการเข้ารหัส

<wsdl:algorithm> คือ ขั้นตอนวิธีการเข้ารหัส

<wsdl:key> คือ ข้อความไซเฟอร์ของกุญแจ

<wsdl:cipher> คือ ข้อความไซเฟอร์ของเอกสาร WSDL

```
<wsdl:cdescription>
<wsdl:algorithm>AES/128/PKCS5Padding</wsdl:algorithm>
<wsdl:key>43d5537e82eca141f2a1b8088298c6a9ef
</wsdl:key><wsdl:cipher>ba54740f5b7813ee51a3e229c8c
23d1cabe24af906e132a442f4ffb4a12f90872225aefc45ce4
c8e47cd6faa079fce40508b59ed9efc295a4c307af65ffc6ca3
0b4f80ba996833f39b25a92122ae22172e47bfc0e99cc04c4
bf8869d95b0074d6176ea9acb840aff1dae2a8fc802f8d0b0e
6</wsdl:cipher></wsdl:cdescription>
```

ภาพที่ 9: ตัวอย่างแท็ก <wsdl:cdescription> ในการเข้ารหัส WSDL

Algorithm2: Decryption data (WSDL_{EK})

1. Extract <wsdl:cdescription> from WSDL_{EK}
2. Extract <wsdl:key> in <wsdl:cdescription> tag
3. Decrypt cipher key with K_p , (OTK)
4. Extract <wsdl:cipher> in <wsdl:cdescription> tag
5. Decrypt cipher msg with OTK (Msg_{DK})
6. Delete <wsdl:cdescription> tag from WSDL_{EK}
7. Extract each Attr in <message>, <portType> and <service> of WSDL_{EK} ($Attr_{EK}$)
8. **For** each $Attr_{EK}$
9. Extract each Attr of Msg_{DK} ($Attr_{DK}$)
10. MatchDoc ($Attr_{DK}$, $Attr_{EK}$)
11. **Endfor**

ภาพที่ 10 : ขั้นตอนวิธีการถอดรหัส

สำหรับการถอดรหัสเอกสาร เมื่อผู้ใช้บริการได้รับเอกสาร WSDL_{EK} จะถอดรหัสเอกสารด้วยอัลกอริทึมดังภาพที่ 10 โดยมีขั้นตอนย่อยที่สำคัญภายใน ส่วน 2 คือ

1. ขั้นตอนการถอดรหัสเอกสาร WSDL (PDecrypt)

สกัดแท็กที่ระบุส่วนเข้ารหัส แล้วจึงถอดรหัสเอกสาร WSDL_{EK} โดยถอดรหัสกุญแจสมมาตร (OTK) ที่ถูกเข้ารหัสด้วยกุญแจส่วนตัวของผู้ให้บริการ และใช้กุญแจ OTK นี้ในการถอดรหัสเอกสาร WSDL_{EK} จะได้ค่าแอททริบิวต์ที่ค้นด้วยสัญลักษณ์ “|”

2. ขั้นตอนการจับคู่ค่าแอททริบิวต์ (MatchDoc)

เมื่อสกัดแต่ละค่าแอททริบิวต์ที่ค้นด้วยสัญลักษณ์ “|” แล้วนำมาเปรียบเทียบกับค่าแอททริบิวต์ใน WSDL_{EK} ดังขั้นตอนวิธีในภาพที่ 11 และภาพที่ 12 แสดงเอกสาร WSDL ที่ถอดรหัสแล้ว

Algorithm3: MatchDoc(Attr_{DK},Attr_{EK})

1. $Attr_{hash} = H(Attr_{DK})$
2. **If** ($Attr_{hash} == Attr_{EK}$)
3. Remove number from $Attr_{DK}$
4. Replace $Attr_{EK}$ of WSDL_{EK} with $Attr_{DK}$
5. **Endif**

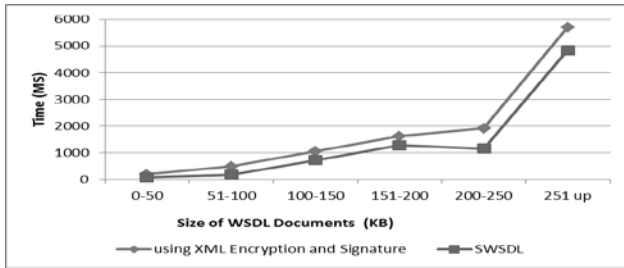
ภาพที่ 11: ขั้นตอนการจับคู่แอททริบิวต์ในเอกสาร WSDL

```
<wsdl:message name="MondkalenderInputMessage">
<wsdl:part name="month" type="s:gMonth"/>
<wsdl:part name="year" type="s:gYear"/>
</wsdl:message>
```

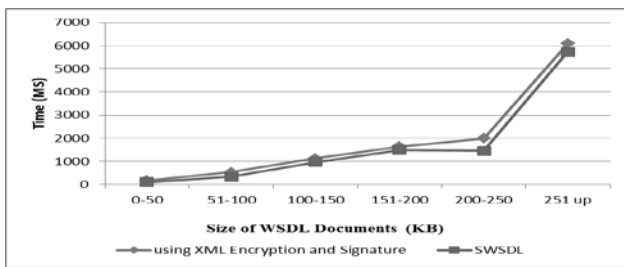
ภาพที่ 12: เอกสาร WSDL ที่ถอดรหัส

4.1 ผลลัพธ์จากการทดลอง

ทดสอบการทำงานโดยนำข้อมูลเอกสาร WSDL จาก wsdllcorpus [13] มาใช้ในการทดสอบ ซึ่งระบบได้ทดสอบกับเอกสาร WSDL 1.1 โดยแบ่งช่วงขนาดของเอกสาร WSDL เป็นช่วงดังนี้คือ 6 1) ขนาด 0- 50 KB 2) ขนาด 51- 100 KB 3) ขนาด 101- 150 KB 4) ขนาด 151 - 200KB 5) ขนาด 201- 250 KB และ 6) ขนาดตั้งแต่ 251 KB ขึ้นไป ผลการทดลองในการเข้าและถอดรหัส WSDL แสดงดังภาพที่ 13 และ 14 ตามลำดับ



ภาพที่ 13 : เวลาเฉลี่ยในการเข้ารหัสเอกสาร WSDL



ภาพที่ 14: เวลาเฉลี่ยในการถอดรหัสเอกสาร WSDL

ผลการทดลองแสดงให้เห็นว่า ระยะเวลาในการเข้ารหัสและถอดรหัสเอกสาร WSDL1.1 ด้วย SWSDL ที่นำเสนอเมื่อเปรียบเทียบกับเข้ารหัสและถอดรหัสด้วยมาตรฐานความปลอดภัยของ XML มีระยะเวลาเฉลี่ยในการเข้ารหัสและถอดรหัสเอกสาร WSDL ลดลง 37.74% และ 22.50 % เมื่อเทียบกับงานวิจัย [11] ตามลำดับ

5. บทสรุป

บทความนี้นำเสนอกลไกความปลอดภัยของเอกสาร WSDL ในเว็บเซอร์วิส ทำให้ผู้ใช้บริการได้รับเอกสาร WSDL ที่มีความลับและมีความถูกต้องสมบูรณ์จากผู้ให้บริการ และช่วยลดระยะเวลาในการเข้ารหัสและถอดรหัสข้อความในเอกสาร WSDL เมื่อเทียบกับงานวิจัย [11] โดยประยุกต์ใช้เอกสาร WSDL1.1 ร่วมกับการเข้ารหัสลับ ด้วยขั้นตอนวิธี AES RSA และกระบวนการแฮช โดยเข้ารหัสเอกสาร WSDL ในส่วนแท็ก <message> แท็ก <portType> และ แท็ก <service> เพื่อป้องกัน WSDL Scanning และ Parameter Tampering โดยงานวิจัยนี้ยังขาดส่วนของกรยืนยันตัวตนของผู้ให้บริการ ซึ่งจะพัฒนาส่วนนี้เพิ่มเติมในอนาคตเพื่อให้ได้กลไกความปลอดภัยที่มีประสิทธิภาพมากขึ้น อีกทั้งจะพัฒนาให้ครอบคลุมเอกสาร WSDL2.0 ด้วย

เอกสารอ้างอิง

- [1] Lucinda Dykes and Ed Tittel, *XML for DUMMIES*, Wiley Publishing: NJ, 2005.
- [2] Kiet T.Tran, *Introduction to Web Services with Java*.
- [3] Manuel Mogollon, *Cryptography and Security Services :Mechanisms and Applications*, Cybertech Publishing: New York, 2007.
- [4] Jothy Rosenberg and David L. Remy, *Securing Web Services with WS-Security*, Sams Publishing: United States of America, 2004.
- [5] W3C, “World Wide Web Consortium”, [Online]. Availablefrom: <http://www.w3.org> [2013,September 21].
- [6] Wikipedia. “Web Services Description Language”, [Online].Availablefrom:http://en.wikipedia.org/wiki/Web_Services_Description_Language [2013,September 20].
- [7] WSDL Threat, “Anatomy of a Web Services attack A Guide to Threats and Preventative Countermeasures”, [Online].Availablefrom:http://csis.org/files/media/csis/insights/anatomy_of_attack_wp.pdf [2013,July 8].
- [8] M.Jensen, N.Gruschka and R.Herkenhoner, “A survey of attacks on web services Classification and Countermeasures”, *Springer*, pp.185-197, 2009.
- [9] N.Sidharth and J.Liu, “ IAPF:A Framework for Enhancing Web Services Security”, *Computer Software and Applications Conference*, vol.1 ,pp. 23-30, 2007.
- [10] A.Mirtalebi and M.R.Khayyambashi, “Enhancing security of Web service against WSDL threats,” *2nd IEEE International Conference*, pp. 920-923, 2011.
- [11] A.Mirtalebi and M.R.Khayyambashi, “A new Security Framework for Protecting WSDL File of Web Service,” *IJCSNS*, vol. 12, no.9, pp. 84-90, 2012.
- [12] N.Shahgholi, M.Mohsenzadeh, M.A.Seyyedi and S.H.Qorani, “A new security framework against Web services' XML attacks in SOA,” *Next Generation Web Services Practices (NWeSP)*, pp.314-319, 2011.
- [13] wsdlcorpus, “WSDL corpus”, [Online]. Availablefrom: <http://coopis.sjtu.edu.cn:8080/cisg/projects/wsdlcorpus/wsdlcorpus.zip/view> [2013,July 7].

สถาปัตยกรรมการกระจายงานของผู้ให้บริการเว็บด้วย NGINX

Load Balancing Architecture for Web Servers with NGINX

สุรชาติ บุญแสน (Surachat Boonsaen)¹ และเนืองวงศ์ ทวยเจริญ (Nuengwong Tuaycharoen)²

สาขาวิศวกรรมคอมพิวเตอร์และโทรคมนาคม

คณะวิศวกรรมศาสตร์ มหาวิทยาลัยธุรกิจบัณฑิตย์

jiw.surachat@gmail.com¹, nuengwong.tum@dpu.ac.th²

บทคัดย่อ

ในปัจจุบัน มีการใช้งานธุรกรรมออนไลน์อย่างแพร่หลาย โดยเฉพาะอย่างยิ่งการขายสินค้าผ่านอินเทอร์เน็ต ทำให้ผู้ใช้บริการเว็บจำเป็นต้องเพิ่มสภาพพร้อมใช้งาน (availability) ของเว็บไซต์ให้สามารถใช้งานได้มากที่สุด บทความนี้ นำเสนอสถาปัตยกรรมการกระจายงานของผู้ให้บริการเว็บที่มีเครื่องแม่ข่ายหลายเครื่อง เพื่อเพิ่มสภาพพร้อมใช้งานในการประมวลผลเว็บไซต์ด้วย NGINX ที่ทำหน้าที่กระจายงาน (Load balance) เพื่อรองรับการใช้งานที่มีจำนวนผู้ใช้งานจำนวนมาก ผลการทดสอบระบบพบว่า สถาปัตยกรรมที่นำเสนอสามารถเพิ่มสภาพพร้อมใช้งานได้ โดยการลดโอกาสที่ระบบจะทำงานไม่ได้ลงเหลือเพียง 0.72 และเมื่อมีผู้ใช้ 1000 คนระบบสามารถให้บริการเว็บได้ภายใน 4 วินาที ซึ่งเป็นเวลาที่ผู้ใช้เว็บยอมรับได้

คำสำคัญ: เทคโนโลยีเว็บ การกระจายงาน ผู้ให้บริการเว็บ NGINX

Abstract

Currently, online transactions are very popular, especially internet shopping. Therefore, websites require more availability to increase their uptime. This article describes the architecture for load balancing of a website with multiple servers to increase the system's availability with NGINX. NGINX is a web server software with Load balancing ability, supporting many users. The experimental results show that the architecture can increase availability by reducing downtime probability to 0.72. And, when a thousand

users simultaneously request services from the website, the architecture's turnaround time is within 4 seconds, which is theoretically acceptable by web users.

Keyword: web technology, load balancing, web server, NGINX

1. บทนำ

ในปัจจุบัน มีการใช้งานธุรกรรมออนไลน์อย่างแพร่หลาย [1] และส่วนหนึ่งของธุรกรรมออนไลน์ที่เป็นที่นิยมอย่างมากคือการขายสินค้าผ่านระบบอินเทอร์เน็ต ที่มีแนวโน้มเพิ่มมากขึ้นเป็นอย่างมาก ซึ่งระบบที่รองรับการขายสินค้าจะประกอบไปด้วย ข้อมูลสินค้า และ รูปภาพ ซึ่งเป็นส่วนสำคัญของการขายที่จะทำให้ผู้ซื้อตัดสินใจเลือกซื้อสินค้านั้น ๆ และที่สำคัญอย่างขาดไม่ได้ คือ ประสิทธิภาพในการส่งข้อมูลมาแสดงผลยังเว็บเบราว์เซอร์ หากไม่มีการจัดการด้านโครงสร้างของเครื่องแม่ข่ายที่ดี ให้รองรับผู้ใช้งานจำนวนมากได้ จะทำให้ผู้ใช้เกิดความรำคาญเมื่อใช้งานไม่ได้ตามต้องการ

จากที่กล่าวมาในข้างต้น ทางเลือกสำหรับการรองรับผู้ใช้งานจำนวนมาก คือ การใช้เครื่องแม่ข่ายหลายเครื่อง แต่หากไม่มีการออกแบบที่ดี เครื่องแม่ข่ายบางเครื่องอาจทำงานหนักเกินไปและบางเครื่องไม่สามารถทำงานได้อย่างเต็มประสิทธิภาพ ซึ่งผลที่ตามมาอาจทำให้ผู้ใช้บริการไม่สามารถใช้งานเว็บไซต์ได้ หรือใช้ได้แต่การประมวลผลของเว็บไซต์ช้า และอาจส่งผลให้เกิดความเสียหายต่อเครื่องแม่ข่ายและกระทบโดยตรงกับผลประกอบการ หากงานด้านการขายไม่สามารถทำการขายผ่านระบบออนไลน์ได้

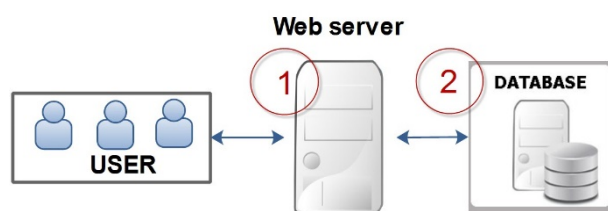
บทความนี้มีวัตถุประสงค์ในการกระจายการทำงานของผู้ให้บริการเว็บเพื่อลดปัญหาการเกิดเหตุการณ์ที่ระบบล่มเหลว

ในรูปของ Downtime ของเว็บไซต์ที่มีสาเหตุมาจากบางส่วนของระบบทำงานหนักเกินไป โดยประยุกต์เข้ากับ NGINX ที่เป็นโปรแกรมแม่ข่ายให้บริการเว็บ เพื่อให้เว็บไซต์สามารถรองรับการทำงานที่หนักและยังสามารถประมวลผลเว็บไซต์ได้ในระดับที่น่าพึงพอใจ เนื่องจากโครงสร้างที่ออกแบบไว้มีการกระจายงาน (Load balancing) เพื่อช่วยลดการทำงานบนเครื่องแม่ข่ายเพียงเครื่องเดียว และทำ High Availability (HA) [4] เพื่อช่วยให้เครื่องแม่ข่ายเครื่องอื่นสามารถทำงานทดแทนเครื่องที่ขัดข้องได้ ซึ่งหากนำสถาปัตยกรรมนี้ไปประยุกต์เข้ากับงานเว็บไซต์ที่มีความเสี่ยงต่อการเกิด Downtime สูง ก็จะช่วยลดการเกิด Downtime บนเครื่องแม่ข่ายได้ และเว็บไซต์ก็สามารถทำงานได้อย่างต่อเนื่อง ซึ่งสถาปัตยกรรมนี้จะรองรับการทำงานร่วมกับ Cloud Server ที่อยู่บนพื้นฐานของเทคโนโลยี Virtualization ที่ทำให้ระบบสามารถมีเครื่องแม่ข่ายได้หลายๆ ที่ โดยใช้ทรัพยากรไม่มากอีกด้วย

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 สถาปัตยกรรมแบบ 3-tier

สถาปัตยกรรม 3-tier เป็นสถาปัตยกรรมผู้ให้บริการเว็บที่ใช้งานกันทั่วไป ที่มีเครื่องแม่ข่ายเว็บเพียงเครื่องเดียว ทำงานกับเครื่องแม่ข่ายฐานข้อมูลเครื่องเดียว ดังแสดงในภาพที่ 1



ภาพที่ 1 : สถาปัตยกรรมแบบ 3-tier

เนื่องจากปัญหาความล้มเหลวที่จุดเดียวของสถาปัตยกรรมแบบนี้ บทความนี้จึงได้ออกแบบสถาปัตยกรรมให้สามารถกระจายงานไปยังเครื่องแม่ข่ายหลายเครื่อง และสามารถทำ Load balance และ HA ได้

2.2 NGINX

NGINX (Engine x) [2] เป็นโปรแกรมให้บริการแม่ข่ายเว็บที่ไม่มีค่าใช้จ่าย ทำงานด้วย Protocol HTTP และมีประสิทธิภาพในการจัดการ reverse proxy รวมไปถึงการ

จัดการเรื่อง Load balancing และ IMAP/POP3 Proxy server โดย NGINX สามารถทำงานได้ทั้งบนระบบปฏิบัติการ Windows และ Linux ซึ่งจะทำหน้าที่เป็นทั้งผู้ให้บริการเว็บและ Load Balance คอยจัดการคำร้องขอบริการเพื่อให้เครื่องแม่ข่ายอยู่ในสถานะการใช้งานที่เป็นปกติ

2.3 Load balancing

Load Balancing[3] คือ การแบ่งจำนวนการทำงานรวมที่ต้องทำระหว่างคอมพิวเตอร์ที่มีจำนวนมากกว่าสองเครื่อง ซึ่งวัตถุประสงค์ของการทำงานในลักษณะนี้เพื่อกระจายการทำงานหรือช่วยกันทำงานของเครื่องแม่ข่ายมากกว่าหนึ่งเครื่อง สามารถจำแนกประเภทการทำงานของ Load balance ออกเป็น 3 ลักษณะได้แก่

- 1) Round-robin เป็นการส่ง traffic ไปยังเครื่องแม่ข่ายในลักษณะของการวนเปลี่ยนกันไปเรื่อยๆ
- 2) Sticky เป็นลักษณะการทำงานที่เป็นการส่ง traffic โดยยึดติดกับ Session ที่ผู้ใช้งานเคยเข้าไปใช้
- 3) Workload เป็นลักษณะการทำงานที่เป็นการส่ง traffic โดยดูที่ปริมาณงานของเครื่องแม่ข่ายในกลุ่มของตนเท่านั้น

2.4 งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องกับการนำ NGINX มาแก้ไขปัญหให้กับเครื่องแม่ข่ายให้บริการเว็บได้แก่ งานวิจัย [6-7] โดยงานวิจัย[6] เน้นการพัฒนาการแก้ปัญหาความแออัดของเครือข่ายด้วยการทำ Web Load Balance และการทำ Cache ให้กับเครื่องแม่ข่ายเว็บ โดยใช้โมดูลของ NGINX ในการทดลองจะนำเครื่องคอมพิวเตอร์ จัดเป็นกลุ่ม กลุ่มละ 3 เครื่อง เครื่องแรกเป็นตัวทำ Load Balance เครื่องที่ 2 และ 3 เป็นแม่ข่ายผู้ให้บริการเว็บ โดยมีติดตั้งระบบ(configuration)ให้ทำ Load Balance แบบ Round-Robin แต่งานวิจัยดังกล่าวยังไม่มีการพิจารณาในเรื่องของ Downtime ที่อาจเกิดขึ้นในตัวโครงสร้าง ดังนั้นเราจึงได้มีการออกแบบและพัฒนาระบบขึ้นเพื่อแก้ปัญหาในเรื่องของการเกิด Downtime ด้วย

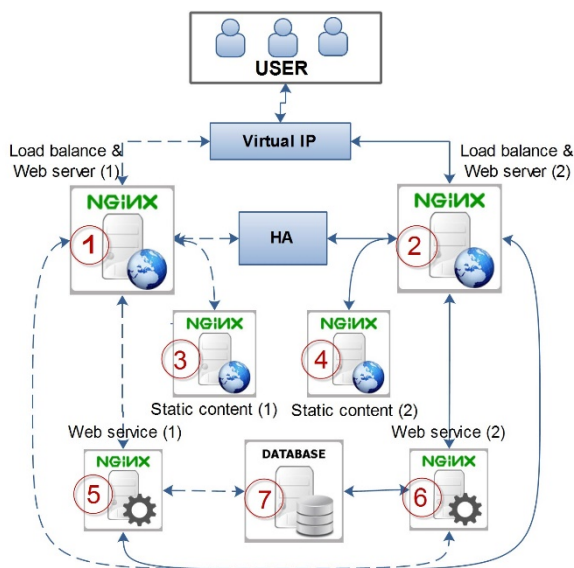
ส่วนในงานวิจัย[7] กล่าวถึงการนำ NGINX มาประยุกต์ใช้ในการแก้ปัญหาเรื่อง High peak load ที่เกี่ยวข้องกับงานด้าน HTTP Streaming ซึ่งวิธีการแก้ปัญหาจะใช้เรื่อง Upstream ของ NGINX ทำเป็น Cache server กั้นระหว่างเครื่องผู้ใช้กับเครื่องแม่ข่ายหลัก เพื่อลดการร้องขอข้อมูลที่ซ้ำๆ และลดภาระงานที่จะเกิดขึ้นในเครื่องแม่ข่าย ซึ่งในบทความ

นี้ได้พัฒนาระบบโดยนำ NGINX มาก่อนก่อนที่การร้องขอจะไปถึงส่วนต่างๆที่อยู่ภายในโครงสร้างหลักเหมือนงานวิจัยนี้ แต่ระบบที่พัฒนาขึ้นมาไปใช้ในลักษณะของการทำ Load Balance

3 รายละเอียดการพัฒนา

3.1 ภาพรวมของระบบ

บทความนี้นำเสนอการพัฒนาสถาปัตยกรรมการกระจายการทำงานของผู้ให้บริการเว็บไปยังเครื่องแม่ข่ายหลายเครื่องเพื่อลดภาระงานของแต่ละส่วน ดังแสดงในภาพที่ 2 ซึ่งสามารถแจกแจงขั้นตอนการทำงานได้ดังนี้



ภาพที่ 2 : สถาปัตยกรรมการกระจายงานของผู้ให้บริการเว็บด้วย NGINX

1) เมื่อการร้องขอบริการของผู้ใช้งานเข้ามา การร้องขอจะวิ่งไปที่ Virtual IP ซึ่งจะใช้ในการทำ HA ของตัว Load balance

2) ตัว Load balance จะคอยจัดการว่า Traffic หรือ URL ที่ผู้ใช้ร้องขอมาเป็นลักษณะใด (ในทางโปรแกรมจะต้องเขียนโปรแกรมแยกด้วยว่า ที่เรียกมาเป็นเว็บเซอร์วิส เป็น static content หรือเป็นการร้องขออย่างอื่น ที่ผู้ให้บริการเว็บจะสามารถรองรับได้ ซึ่งในบทความนี้จะไม่กล่าวถึง)

3) ในการเชื่อมต่อฐานข้อมูลตัวของผู้ให้บริการเว็บ จะต้องผ่านตัวเว็บเซอร์วิสเสมอ เพื่อลดงานของเครื่องแม่ข่ายเว็บ

4) รูปภาพบนเว็บจะถูกประมวลผลโดยเครื่องแม่ข่ายประเภท Static content

5) เครื่องแม่ข่ายของเว็บเซอร์วิส จะมีหน้าที่ประมวลผลให้กับอุปกรณ์อื่น ๆ ด้วย เช่น โทรศัพท์มือถือหรือแท็บเล็ต

6) NGINX จะเป็นตัวจัดการ traffic เพื่อทำ Load balance ให้เครื่องแม่ข่ายประเภทเดียวกัน เช่น เครื่องแม่ข่ายของเว็บเซอร์วิส ก็ทำงานให้เฉพาะการร้องขอเว็บเซอร์วิส

7) กรณีที่เครื่องแม่ข่ายเครื่องใดที่เป็นที่เป็นประเภทเดียวกันมีเหตุขัดข้องที่ไม่สามารถทำงานได้ เครื่องแม่ข่ายประเภทเดียวกันอีกตัวหนึ่งจะทำงานแทนทันที

จากสถาปัตยกรรมในภาพที่ 2 ผู้วิจัยได้แบ่งประเภทของการให้บริการของเครื่องแม่ข่าย เพื่อให้เครื่องแม่ข่ายแต่ละเครื่องสามารถแบ่งงานกันได้อย่างเหมาะสมตามประเภทของไฟล์ ดังแสดงในตารางที่ 1 จะเห็นได้ว่าการใช้เครื่องแม่ข่ายมากกว่าหนึ่งเครื่องสำหรับแม่ข่ายแต่ละประเภท ยกเว้นเครื่องแม่ข่าย Database ซึ่งใช้เพียงเครื่องเดียวเพื่อให้สามารถเปรียบเทียบกับเครื่องแม่ข่ายเว็บแบบ 3-tier ทั่วไปได้ การแบ่งเครื่องแม่ข่ายออกเป็น web server และ web service เพื่อแยกเครื่องแม่ข่ายสำหรับการร้องขอจากเว็บและอุปกรณ์เคลื่อนที่ออกจากกัน ถึงแม้ว่าจะเชื่อมต่อกันแบบอนุกรม แต่ web server ทั้งสองสามารถเรียกใช้งาน web service ตัวใดก็ได้ ส่วน static content ถูกแยกไปประมวลผลต่างหาก ทำให้ถ้าเครื่องแม่ข่าย static content เสียหาย ระบบเว็บยังสามารถทำงานได้ ถึงแม้ว่าจะแสดงผลไม่ครบถ้วนก็ตาม

ตารางที่ 1: ตัวอย่างประเภทไฟล์ที่แยกตามเครื่องแม่ข่าย

เครื่องแม่ข่าย	ประเภทไฟล์ที่จัดการ
Web server (1)	*.php, *.html, *.jsp
Web server (2)	*.php, *.html, *.jsp
Static content (3)	*.jpg, *.png, *.css, *.js
Static content (4)	*.jpg, *.png, *.css, *.js
Web service (5)	*.php, *.jsp
Web service (6)	*.php, *.jsp
Database(7)	*.sql

3.2 กระบวนการทำ Load Balance

ภาพที่ 3 แสดงตัวอย่างรหัสในไฟล์ nginx.conf ที่เป็นการทำงานหลักในการทำ Load Balance

```

1 upstream web {
2     server 192.168.1.11
3         max_fails=3
4         fail_timeout=10s;
5     server 192.168.1.12
6         max_fails=3
7         fail_timeout=10s;
8 }
9
10 server {
11     listen 80;
12     server_name www.test.local;
13     location / {
14         proxy_pass http://web;
15     }
16 }

```

ภาพที่ 3: ตัวอย่างรหัสตั้งค่าระบบของ NGINX

จากสถาปัตยกรรมในภาพที่ 2 กำหนดให้ Virtual IP คือ 192.168.1.10 และภาพที่ 3 บรรทัดที่ 2 และ 5 กำหนดให้เครื่องแม่ข่ายเว็บหมายเลข 1 และ 2 ตั้งค่า IP เป็น 192.168.1.11 และ 192.168.1.12 ตามลำดับ ทำ Load Balance กันโดยใช้ upstream ที่มีตัวแปรชื่อ web (บรรทัดที่ 14) และหากมีการร้องขอบริการด้วย www.test.local ที่ port 80 (บรรทัดที่ 11-12) การร้องขอจะถูกส่งไปที่เครื่องแม่ข่ายเว็บทั้ง 2 เครื่องในเซตของ upstream (บรรทัดที่ 1-2 และ 5) ซึ่งเป็นการทำ Load Balance แบบ Round-Robin

3.3 กระบวนการทำ High Availability (HA)

1) HA ระหว่าง Load Balance เครื่องแม่ข่ายหมายเลข 1 และ 2 ผู้วิจัยได้ออกแบบโครงสร้างให้เครื่องแม่ข่ายทำงานที่ละตัวในลักษณะของ Cluster โดยใช้โปรแกรมชื่อ Heartbeat ซึ่งสามารถแสดงตัวอย่างรหัสตั้งค่าระบบในไฟล์ ha.cf ได้ดังภาพที่ 4

2) HA ระหว่าง Web Server, Static Content และ Web Service ในส่วนนี้จะเป็นการทำ HA ของเครื่องแม่ข่ายที่อยู่ภายใต้ Load Balance ซึ่งจะใช้โมดูลของ NGINX ช่วยในการตรวจสอบว่าเครื่องแม่ข่ายที่มีการร้องขอบริการไป สามารถทำงานได้หรือไม่ ถ้าไม่สามารถทำงานได้ การร้องขอจะถูกส่งไปทำงานยังเครื่องที่ทำงานได้อยู่ โดยสามารถแสดงรหัส

ตั้งค่าระบบได้ดังภาพที่ 3 บรรทัดที่ 3-4 และ 6-7 โดยระบุ max_fails = 3 ซึ่งเป็นค่าจำนวนครั้งที่ระบบยอมให้เชื่อมต่อกับเครื่องแม่ข่ายไม่สำเร็จ และ fail_timeout = 10s เป็นการระบุว่าเกิด time out ได้ 10 วินาที

```

1 logfile /var/log/ha-log
2
3 autojoin none
4 auto_failback off
5 bcast eth3
6 warntime 5
7 deadtime 15
8 initdead 60
9 keepalive 2
10
11 node lb1 #---> (เซิร์ฟเวอร์หมายเลข 1 )
12 node lb2 #---> (เซิร์ฟเวอร์หมายเลข 2 )

```

ภาพที่ 4: แสดงตัวอย่างรหัสตั้งค่าระบบในไฟล์ ha.cf

4. การทดสอบการใช้งาน

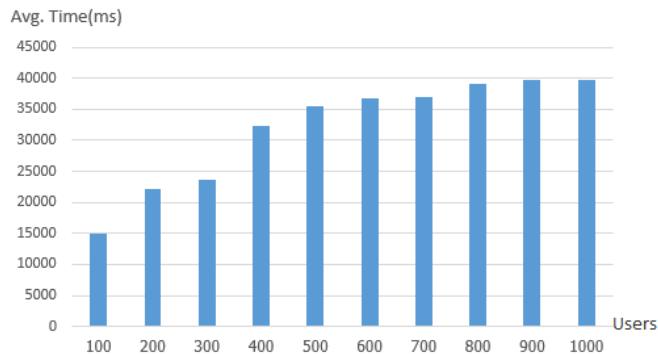
4.1 สภาพแวดล้อมในการทดสอบ

ผู้วิจัยได้ทดสอบระบบกับข้อมูลเว็บเพจแสดงสินค้าจำนวน 100 รายการจากเว็บไซต์จัดจำหน่ายสินค้าแห่งหนึ่ง[8] โดยใช้เครื่องมือชื่อ Apache Jmeter [2] ในการจำลองการร้องขอบริการของผู้ใช้ และวัดประสิทธิภาพทางเวลาของการดาวน์โหลดหน้าเว็บ โดยกำหนดช่วงผู้ใช้งานตั้งแต่ 100 คนเพิ่มขึ้นจนถึง 1000 คน และยังมีปัจจัยอื่นๆที่เกี่ยวข้อง ดังนี้

- 1) อินเทอร์เน็ตความเร็วเฉลี่ย 7.0 Mbps
- 2) ขนาดข้อมูลของหน้าเว็บไซต์ 3.09 MB
- 3) Traffic ที่อยู่ในเครือข่าย มีการใช้งานเฉพาะการทดสอบนี้เท่านั้น
- 4) ติดตั้ง Linux CentOS 6.3 และติดตั้ง NGINX เป็นผู้ให้บริการเว็บที่เครื่องหมายเลข 1 ถึง 6 และ ติดตั้ง MySQL เพื่อให้บริการ Database ที่เครื่องหมายเลข 7
- 5) ตั้งค่า IP ที่เครื่องหมายเลข 1 ถึง 7 เป็น 192.168.1.[11-17]/24 ตามลำดับ
- 6) สร้าง sub interface ที่เครื่องแม่ข่ายหมายเลข 1 ให้ IP เป็น 192.168.1.10 ซึ่งจะใช้เป็น Virtual IP สำหรับทำ Cluster ที่ เครื่องหมายเลข 1 และ 2
- 7) ติดตั้ง Heartbeat ที่เครื่องหมายเลข 1 และ 2

4.2 ผลการทดสอบและการวิจารณ์ผล

จากการทดสอบประสิทธิภาพความเร็วในการดาวน์โหลดหน้าเว็บ เวลาที่ใช้ในการดาวน์โหลดตามการเพิ่มจำนวนผู้ใช้งานสามารถแสดงได้ดังภาพที่ 5



ภาพที่ 5 : กราฟแสดงจำนวนผู้ใช้ในเวลาเดียวกันและเวลาที่ใช้ในการดาวน์โหลดเว็บไซต์

จากภาพที่ 5 พบว่าสถาปัตยกรรมที่ออกแบบขึ้นสามารถรองรับจำนวนผู้ใช้ที่เพิ่มขึ้น โดยเวลาที่ใช้ในการดาวน์โหลดหน้าเว็บไซต์เพิ่มขึ้นไม่มากนัก และเมื่อจำนวนการร้องขอสูงถึง 1,000 คน ระบบก็ยังสามารถรองรับผู้ใช้ให้ดาวน์โหลดเว็บได้ภายในเวลาเพียง 4 วินาที ซึ่งเป็นเวลาที่ผู้ใช้เว็บทั่วไปยอมรับได้ [9]

นอกจากการทดสอบด้านประสิทธิภาพในการดาวน์โหลดแล้ว ยังมีการทดสอบเรื่อง Downtime ที่จะเกิดขึ้นเมื่อเกิดความล้มเหลวที่เครื่องแม่ข่ายตัวใดตัวหนึ่งในกลุ่มที่ให้บริการ ซึ่งผลการทดสอบของสถาปัตยกรรมแบบ 3-tier ทั่วไปแสดงในตารางที่ 2 และผลการทดสอบของสถาปัตยกรรมที่พัฒนาขึ้นแสดงในตารางที่ 3 โดยกำหนด No.1-7 เป็นหมายเลขเครื่องแม่ข่ายแต่ละเครื่องในภาพที่ 1 และ 2 และสัญลักษณ์ต่าง ๆ มีความหมายดังนี้

- ✓ คือ สถานะปกติ,
- ✗ คือ ไม่ปกติ (จำลองโดยการปิดการทำงาน)
- D คือ สถานะปกติหรือไม่ปกติ(ปิดอยู่)ก็ได้ (Don't care)

จากผลการทดสอบในตารางที่ 2 พบว่า ระบบแบบสถาปัตยกรรม 3-tier มีโอกาสเกิด Downtime เป็น 0.75 (3 จาก 4 กรณี)

ตารางที่ 2 : ผลการทดสอบระบบแบบสถาปัตยกรรม 3-tier

กรณีทดสอบ	No.1	No.2	สถานะระบบ
1	✓	✓	✓
2	✓	✗	✗
3	✗	✓	✗
4	✗	✗	✗

ตารางที่ 3 : ผลการทดสอบสถาปัตยกรรมที่พัฒนาขึ้นเมื่อปิดการทำงานของเครื่องแม่ข่ายบางเครื่อง

กรณีทดสอบ	No. 1	No. 2	No. 3	No. 4	No. 5	No. 6	No. 7	สถานะระบบ
1	✗	✓	✓	✓	✓	✓	✓	✓
2	✓	✗	✓	✓	✓	✓	✓	✓
3	✓	✓	✗	✓	✓	✓	✓	✓
4	✓	✓	✓	✗	✓	✓	✓	✓
5	✓	✓	✓	✓	✗	✓	✓	✓
6	✓	✓	✓	✓	✓	✗	✓	✓
7	D	D	D	D	D	D	✗	✗
8	✗	✗	D	D	D	D	✓	✗
9	เครื่องใดเครื่องหนึ่งทำงานได้		✗	✗	เครื่องใดเครื่องหนึ่งทำงานได้		✓	✓
10	เครื่องใดเครื่องหนึ่งทำงานได้		D	D	✗	✗	✓	✗

ส่วนระบบที่พัฒนาขึ้นมีผลการทดสอบดังแสดงในตารางที่ 3 ซึ่งโอกาสที่ระบบที่พัฒนาขึ้นจะล้มเหลวเป็นไปได้เมื่อเกิด 3 เหตุการณ์ คือเหตุการณ์กรณีทดสอบที่ 7, 8, และ 10 เท่านั้น ซึ่งสามารถคำนวณโอกาสที่จะเกิด Downtime ของระบบที่พัฒนาขึ้นตามทฤษฎีความน่าจะเป็นได้ดังนี้

กรณีที่เป็นไปได้ทั้งหมด: $2^7 = 128$ กรณี

ทั้งระบบล้มเหลวเพราะกรณีทดสอบที่ 7: $2^6 = 64$ กรณี

ทั้งระบบล้มเหลวเพราะกรณีทดสอบที่ 8: $2^4 = 16$ กรณี

ทั้งระบบล้มเหลวเพราะกรณีทดสอบที่ 10: $3 \times 2^2 = 12$ กรณี
 ดังนั้นโอกาสที่ระบบที่จะล้มเหลว $= (64+16+12)/128 = 0.72$
 จะเห็นว่า ระบบที่พัฒนาขึ้นมีโอกาสเกิด Downtime (0.72)
 น้อยกว่าระบบแบบสถาปัตยกรรมแบบ 3-tier (0.75)

5. บทสรุป

เนื่องจากผู้ให้บริการเว็บทั่วไปมักใช้สถาปัตยกรรมแบบ 3-tier ในการให้บริการเว็บไซต์ที่มีทั้งข้อความ ภาพ และฐานข้อมูล ซึ่งเสี่ยงต่อการล้มเหลวของระบบ ผู้วิจัยจึงได้พัฒนาสถาปัตยกรรมผู้ให้บริการเว็บที่สามารถกระจายงานให้กับเครื่องแม่ข่ายหลายเครื่องได้ โดยใช้ NGINX เป็นโปรแกรมกระจายงาน และแบ่งเครื่องแม่ข่ายตามประเภทของไฟล์ที่ให้บริการ จากผลการทดสอบระบบสรุปได้ว่าสถาปัตยกรรมการกระจายงานเพื่อการประมวลเว็บไซต์ที่น่าเสนอในบทความนี้มีประสิทธิภาพที่ยอมรับได้โดยสามารถรองรับผู้ใช้จำนวน 1,000 คนได้ภายในเวลา 4 วินาที และมีโอกาสเกิด downtime น้อยกว่าสถาปัตยกรรมแบบ 3-tier ที่ทำงานด้วยเครื่องแม่ข่ายเดียว

เอกสารอ้างอิง

- [1] ข้อมูลอินเทอร์เน็ตแบนด์วิดท์.[ออนไลน์] เข้าถึงได้จาก : <http://internet.nectec.or.th/webstats/bandwidth.iir?Sec=bandwidth>. สืบค้น 10 ธันวาคม 2556
- [2] NGINX.[ออนไลน์] เข้าถึงได้จาก : <http://wiki.NGINX.org/>. สืบค้น 22 ธันวาคม 2556
- [3] Load balancing.[ออนไลน์] เข้าถึงได้จาก : <http://serverlab.ca/tutorials/linux/web-services/load-balancing-web-servers-with-NGINX-on-centos-6>. สืบค้น 24 ธันวาคม 2556
- [4] High Availability.[ออนไลน์] เข้าถึงได้จาก : <http://www.linux-ha.org/wiki/>. สืบค้น 25 ธันวาคม 2556
- [5] Apache Jmeter.[ออนไลน์] เข้าถึงได้จาก : <https://jmeter.apache.org/>
- [6] Xiaoni Chi; Bichuan Liu; Qi Niu; Qiuxuan Wu, "Web Load Balance and Cache Optimization Design Based Nginx under High-Concurrency Environment," Digital Manufacturing and Automation (ICDMA), 2012 Third International Conference on , pp.1029,1032, July 31 2012-Aug. 2 2012.

[7] Weikai Xie; Yunfei Li; Chenping Lu; Ruimin Shen, "Optimizing the resource-updating period behavior of HTTP cache servers for better scalability of live HTTP streaming systems," Broadband Multimedia Systems and Broadcasting (BMSB), 2012 IEEE International Symposium on, pp.1,6, 27-29 June 2012.

[8] บริษัท ออฟฟิศเมท จำกัด มหาชน. ข้อมูลสินค้าที่ใช้ในการ

ทดสอบ.[ออนไลน์] เข้าถึงได้จาก : <http://www.officemate.co.th>

[9] Kissmetrics, "How loading time affects your bottom line". <http://goo.gl/kosva>

การเปรียบเทียบวิธีการเลือกคุณลักษณะที่เหมาะสมเพื่อการจัดหมวดหมู่เว็บเพจผิดกฎหมาย โดยใช้เทคนิคการทำเหมืองข้อมูล

Comparison of Feature Selection Methods for Inappropriate Webpage Classification by Data Mining Technique

นำทิพย์ มากนคร (Namthip Marknakorn)¹ และมาลีรัตน์ โสदानิล (Maleerat Sodanil)²

^{1,2}ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
je-ab-t@hotmail.com¹, msn@kmutnb.ac.th²

บทคัดย่อ

การทำเหมืองข้อมูลเว็บเพจซึ่งเป็นข้อมูลที่มีขนาดใหญ่ อาจประสบปัญหาต่าง ๆ เช่น การใช้เวลานานในการสร้างโมเดล มิติของข้อมูลที่มีจำนวนมาก การกระจายของข้อมูล เป็นต้น งานวิจัยนี้จึงได้นำเสนอการเปรียบเทียบวิธีการเลือกคุณลักษณะที่เหมาะสม ด้วยวิธีการจัดอันดับแบบ Information Gain, Gain Ratio และ Chi-Square โดยใช้เทคนิคซัพพอร์ตเวกเตอร์แมชชีนเนลฟังก์ชันแบบเรเดียลเบสิสฟังก์ชันในการจัดหมวดหมู่เว็บเพจ จำนวน 5,399 URL ผลการทดสอบพบว่าวิธีการจัดอันดับแบบ Chi-Square ให้ผลลัพธ์ที่มีประสิทธิภาพดีที่สุดกล่าวคือ เมื่อลดจำนวนแอตทริบิวต์ 50% ให้ค่าความถูกต้อง (Accuracy) ร้อยละ 95.98 และเมื่อทำการปรับค่าพารามิเตอร์ Gamma ส่งผลให้ค่าความถูกต้องเพิ่มขึ้นโดยการคัดเลือกคุณลักษณะด้วยวิธีการจัดอันดับแบบ Chi-Square ให้ค่าความถูกต้องเพิ่มขึ้นจาก ร้อยละ 95.94 เป็น ร้อยละ 97.93 ที่การลดจำนวนแอตทริบิวต์ 30% และการปรับค่าพารามิเตอร์ Gamma เท่ากับ 1.1 แสดงให้เห็นว่าการคัดเลือกคุณลักษณะและการปรับค่าพารามิเตอร์ของเทคนิคซัพพอร์ตเวกเตอร์แมชชีนเนลฟังก์ชันแบบ RBF สามารถเพิ่มประสิทธิภาพการจัดหมวดหมู่เว็บเพจผิดกฎหมายได้

คำสำคัญ: การคัดเลือกคุณลักษณะ ซัพพอร์ตเวกเตอร์แมชชีน
การจัดหมวดหมู่เว็บเพจ

Abstract

Web data mining on large data have various problems such as used long time to build models , many

dimensional data etc. This paper is to make the comparison of feature selection methods for inappropriate webpage classification using Information Gain , Gain Ratio and Chi-Square with used SVM-RBF to classify the 5,399 are collected from web. The experimental show that Chi-Square is the best method when reduce 50% of attributes the accuracy given 95.98% . The Chi-Square gave the accuracy increased from 95.94% to 97.93 at reducing 30% of attributes and the adjusted gamma parameter is 1.1 that shown classification and the adjusted parameter of RBF kernel technique be able to improve the performance in terms of accuracy.

Keyword : Feature Selection, Support Vector Machine, Web Document Classification

1. บทนำ

เทคโนโลยีการสื่อสารที่พัฒนาอย่างรวดเร็ว ทำให้เกิดการจับเก็บและการเผยแพร่ข้อมูลในรูปแบบข้อมูลอิเล็กทรอนิกส์ ซึ่งหนึ่งในข้อมูลที่เป็นรูปแบบอิเล็กทรอนิกส์คือ ข้อมูลประเภทเอกสารเช่น เว็บเพจ (Web page) การสร้างเว็บเพจเป็นเรื่องที่ง่ายและสะดวกรวดเร็ว ทำให้ข้อมูลเว็บเพจมีปริมาณมากขึ้น และมีความหลากหลายทำให้ยากต่อการค้นหาและการจัดหมวดหมู่เอกสาร ดังนั้นเพื่อให้การจับเก็บและสืบค้นเอกสารง่ายและเป็นไปตามความต้องการจึงจำเป็นต้องใช้วิธีการจัดหมวดหมู่ให้สอดคล้องและตรงกับดัชนี จึงได้มีการศึกษา

เกี่ยวกับการนำวิธีการเรียนรู้ด้วยคอมพิวเตอร์มาประยุกต์ใช้ในการประมวลผลภาษาธรรมชาติ เพื่อจัดหมวดหมู่เอกสาร ได้แก่ การทำเหมืองข้อมูล (Data Mining) โดยวิธีการจัดหมวดหมู่ (Classification) ซึ่งจะเกี่ยวข้องกับการวิเคราะห์ข้อมูลที่มีจำนวนมหาศาล โดยประกอบไปด้วย คุณลักษณะที่ไม่ตรงประเด็น และมีมิติของข้อมูลจำนวนมากส่งผลให้การทำเหมืองข้อมูลต้องใช้เวลาในการวิเคราะห์มากขึ้น ถ้าข้อมูลมีมิติ หรือตัวแปรมากจะทำให้ข้อมูลเกิดการกระจาย และอาจไม่มีความสัมพันธ์กับมิติอื่น [1]

ดังนั้นการคัดเลือกคุณลักษณะเพื่อลดมิติของข้อมูล จึงเป็นแนวทางหนึ่งที่สามารถช่วยแก้ปัญหาดังกล่าว และช่วยให้การจัดหมวดหมู่ข้อมูลได้แม่นยำมากขึ้น เพื่อค้นหาวิธีการคัดเลือกคุณลักษณะที่เหมาะสม ผู้วิจัยจึงได้ทำการศึกษาเปรียบเทียบประสิทธิภาพของเทคนิคการคัดเลือกคุณลักษณะด้วยวิธีการจัดอันดับ โดยใช้เทคนิคซอฟต์แวร์เคอร์แมชชีนสำหรับการจัดหมวดหมู่

2. ทฤษฎีและวรรณกรรมที่เกี่ยวข้อง

2.1 การคัดเลือกคุณลักษณะ

การคัดเลือกคุณลักษณะ คือการลดขนาดของข้อมูลโดยการทำให้อัตราส่วนข้อมูลเดิมมีขนาดลดลงและสูญเสียลักษณะสำคัญของข้อมูลน้อยที่สุด เทคนิคการคัดเลือกที่ต่างกันทำให้ได้คุณลักษณะที่ต่างกันด้วย การคัดเลือกคุณลักษณะ (Feature Selection) มีชื่อเรียกหลายอย่าง ในงานด้านสถิติมักเรียกว่า variable selection เพราะมองว่าคุณลักษณะแต่ละอันคือตัวแปรแบบสุ่ม (random variable) บางทีเรียกว่า subset selection เพราะการเลือกของจำนวนหนึ่งออกจากของทั้งหมดก็คือการเลือก subset นั้นเอง ในงานวิจัยนี้เลือกวิธีการคัดเลือกคุณลักษณะมาเพื่อเปรียบเทียบ 3 วิธี ได้แก่

2.1.1 Information Gain

เป็นเทคนิคการคัดเลือกมิติของข้อมูล โดยการจัดอันดับ ซึ่งมีการคำนวณค่า Gain ที่มากที่สุดของแต่ละมิติข้อมูล เพื่อจะถูกเลือกให้เป็นกลุ่มย่อยที่มีความสามารถในการจำแนกข้อมูล[2] ตามสมการ(1)(2)

$$Entropy = -\sum_t p(j|t) \log_2 p(j|t) \quad (1)$$

โดยที่ $-\sum_t$ คือ ผลรวมของความน่าจะเป็นของ j ที่เกิดในคลาส t

$$Gain = Entropy(p) - \left[\sum_{i=1}^k \frac{n_i}{n} Entropy(i) \right] \quad (2)$$

โดยที่ Entropy(p) คือค่าเอนโทรปีของค่าราก (Root)

$\left[\sum_{i=1}^k \frac{n_i}{n} Entropy(i) \right]$ คือ ค่าเอนโทรปีของแต่ละโหนดย่อย

2.1.2 Gain Ratio

เป็นวิธีการประเมินความน่าเชื่อถือของมิติของข้อมูล โดยการวัดค่า Gain Ratio ในแต่ละคลาสการคำนวณ GR โดยใช้ค่า SplitINFO ในสมการที่ (3) และการคำนวณค่าการวัด Gain Ratio [3]ดังสมการที่ (4)

$$SplitINFO = -\sum_{i=1}^k \frac{n_i}{n} \log_2 \frac{n_i}{n} \quad (3)$$

$$GainRatio = \frac{\Delta INFO}{SplitINFO} \quad (4)$$

2.1.3 Chi Squared (x^2)

การประเมินค่าของ แอททริบิวต์โดยคำนวณค่า Chi-Square ทางสถิติ [3] ดังแสดงใน สมการที่ (5)

$$x^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} \quad (5)$$

2.2 การจัดหมวดหมู่เอกสารประเภทเว็บเพจ

การจัดหมวดหมู่เอกสารประเภทเว็บเพจ (Web Document Classification) มีความแตกต่างจากการจัดหมวดหมู่เอกสารทั่วไป (Text Categorization) กล่าวคือ เอกสารทั่วไปมักมีโครงสร้าง เนื้อหา และคุณลักษณะที่ชัดเจนซึ่งแตกต่างจากเว็บเพจเนื่องจากเว็บเพจประกอบด้วยตัวอักษร รูปภาพ เสียง หรือภาพวิดีโอ ข้อมูลบนเว็บเพจถูกกำกับด้วยแท็ก ตามมาตรฐานของเอกสารไฮเปอร์เท็กซ์

การจัดหมวดหมู่เอกสารประเภทเว็บเพจใช้พื้นฐานของการทำนาย (Prediction) โดยเอกสารถูกติดฉลากด้วยหัวข้อ (Topic) และนำชุดเอกสารเอกสารที่ติดฉลากนี้ไปใช้ในการเรียนรู้ด้วยระบบเพื่อนำไปทำนายเอกสารใหม่ว่าเป็นหัวข้อใด วิธีที่ใช้ในการเรียนรู้ที่ได้รับความนิยมคือ การเรียนรู้แบบมีผู้สอน (Supervised Learning) โดยการเรียนรู้จากชุดข้อมูลที่ติดฉลากด้วยคลาส (Class) หรือหมวดหมู่ (Category) ซึ่งชุด

ข้อมูลนี้ใช้เป็นชุดฝึกฝนให้กับระบบและนำไปใช้ทำนายชุดข้อมูลใหม่ต่อไป

2.3 เทคนิคซัพพอร์ตเวกเตอร์แมชชีน

ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) เป็นเทคนิคหนึ่งที่ใช้สำหรับการจำแนกประเภทหรือการจัดหมวดหมู่ข้อมูล (Data Classification) โดยการวิเคราะห์จะเป็นแนวทางของการคำนวณหาเส้นแบ่งกลุ่มที่เป็นเชิงเส้นและไม่เป็นเชิงเส้น โดยพิจารณาจากจุดของข้อมูลที่อยู่ในกลุ่มที่แตกต่างกันจุดเหล่านี้เรียกว่าซัพพอร์ตเวกเตอร์ซึ่งจะนำมาหาสมการเส้นแบ่งกลุ่มที่เหมาะสมต่อไป SVM จะมีทั้งการจำแนกประเภทแบบไบนารี เพื่อสร้างสมการเส้นตรงในการแบ่งเขตข้อมูลออกเป็น 2 กลุ่มและการจำแนกประเภทแบบหลายกลุ่ม (Multiclass SVM) เพื่อแก้ปัญหาการจำแนกประเภทที่มีมากกว่าสองกลุ่ม โครงสร้างของ SVM ประกอบด้วย 2 ส่วนหลัก คือ การเพิ่มระยะการจำแนกมากที่สุด และการลดข้อผิดพลาดให้ต่ำที่สุด สำหรับการวัดความคล้าย (Kernel) ของซัพพอร์ตเวกเตอร์แมชชีนที่ใช้ในงานวิจัยนี้คือ แบบ Radial Basis Function: RBF ซึ่งเป็นการคำนวณหาขอบเขตข้อมูลโดยอาศัยวิธีแบบ Radial Basis เข้ามาช่วยในการคำนวณ ดังแสดงไว้ในสมการที่ (6)

$$\text{emp}(-\text{gamma} \times |u - v|^2) \quad (6)$$

2.4 วิธีการวิเคราะห์ความแม่นยำตรงของโมเดล k-fold cross-validation

การตรวจสอบไขว้กัน (Cross-Validation) [4] เป็นวิธีการตรวจสอบค่าความผิดพลาดในการคาดการณ์ของโมเดล โดยพื้นฐานวิธีการ การตรวจสอบไขว้กัน k-Fold Cross-validation เป็นวิธีการที่แบ่งข้อมูลออกเป็นกลุ่มจำนวน k กลุ่ม (k-Fold) ในตอนแรกเลือกข้อมูลกลุ่มที่ 1 เป็นข้อมูลชุดทดสอบ และข้อมูลชุดที่เหลือจะเป็นข้อมูลชุดสอน นำข้อมูลไป จัดหมวดหมู่ จากนั้นจะสลับข้อมูลกลุ่มที่ 2 มาเป็นชุดทดสอบและข้อมูลกลุ่มอื่นๆที่เหลือเป็นชุดทดสอบ สลับอย่างนี้ไปเรื่อยๆจนครบ k กลุ่ม ในขั้นตอนสุดท้ายจะหาค่าเฉลี่ยของค่าความถูกต้องในแต่ละกลุ่ม วิธีการนี้ข้อมูลทุกตัวอย่างจะได้เป็นทั้งชุดทดสอบและชุดสอน

2.5 งานวิจัยที่เกี่ยวข้อง

การจัดหมวดหมู่เว็บเพจนั้นมีวิธีการจัดหมวดหมู่หลายวิธีที่สามารถนำมาใช้ได้ขึ้นอยู่กับลักษณะของข้อมูลและความต้องการของระบบ

ชัยพร [5] ได้เปรียบเทียบประสิทธิภาพในการจำแนกเว็บเพจ โดยนำมาทำการทดลองร่วมกับอัลกอริธึมการกรองแบบ Information Gain ร่วมกับอัลกอริธึมต่าง ๆ 4 อัลกอริธึมคือ PART, Naïve Bays, Decision Tree และ SVM ซึ่งผลที่ได้แสดงให้เห็นว่าอัลกอริธึม Information Gain ร่วมกับ SVM ให้ความแม่นยำสูงที่สุดในการจำแนกประเภทเว็บเพจ

ศุภเทพ [6] การจำแนกประเภทข้อความสั้นภาษาไทย จากเว็บไซต์ทวิตเตอร์ โดยเลือกอัลกอริธึม 3 อัลกอริธึม คือ Naïve-Bayes Support Vector Machine(SVM) และ Tree ผลการทดสอบประสิทธิภาพปรากฏว่า อัลกอริธึม SVM สามารถจำแนกประเภทได้ถูกต้องมากที่สุด

นิยม [7] กระบวนการการจัดหมวดหมู่เว็บเพจนั้นจะใช้การสกัดคุณลักษณะของเว็บเพจ (Feature Extraction) โดยใช้การสกัดคุณลักษณะที่ปรากฏบนหน้าเว็บเพจ (On-Page Features) ได้แก่คุณลักษณะจากแท็ก Title, Meta และ Body เพื่อพิจารณาถึงความสำคัญของคุณลักษณะในแต่ละส่วนและแทนค่าความ สำคัญนั้นด้วยค่าปัจจัยน้ำหนักที่เหมาะสมที่สุด

จิเรขา [8] การจัดหมวดหมู่เว็บเพจซึ่งใช้ข้อมูลบนเว็บเพจที่มีจำนวนมากทำให้การประมวลผลซ้ำการเพิ่มประสิทธิภาพของการประมวลผลการจำแนกเว็บเพจทำได้โดยการคัดเลือกคุณลักษณะที่เหมาะสม (Feature Selection) และลดคุณลักษณะที่ไม่จำเป็นลงเพื่อให้ผลจากการตัดคำถูกต้อง และนำคำมาจัดหมวดหมู่เอกสาร

วิภาวรรณ [1] การเปรียบเทียบเทคนิคการคัดเลือกคุณลักษณะของการกรอง (Filter Selection) 3 เทคนิคได้แก่ information gain, gain ratio และ chi-square ผลสรุปว่า การคัดเลือกคุณลักษณะแบบการกรองด้วยเทคนิค chi-square ให้ประสิทธิภาพดีที่สุด

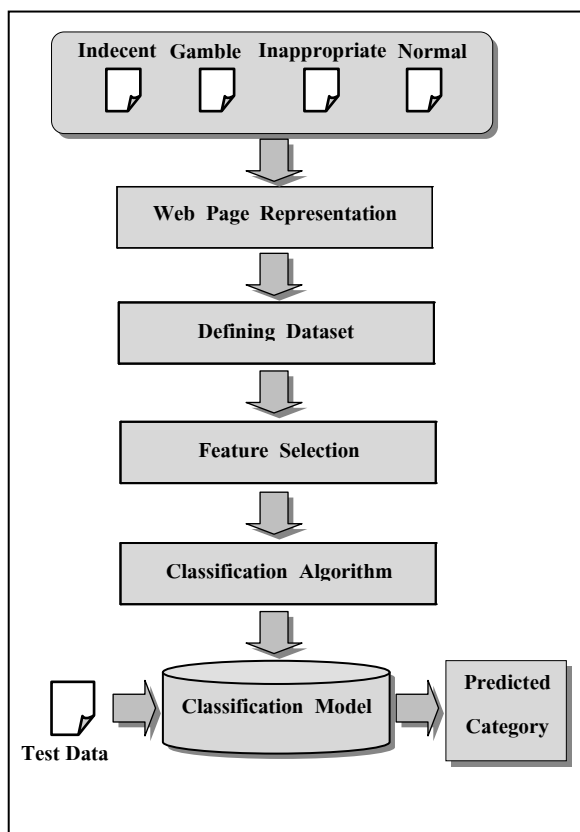
วงกต [10] การเปรียบเทียบประสิทธิภาพในการลดมิติของข้อมูลด้วย information gain, gain ratio และ Linear SVM Weight โดยทดสอบกับข้อมูลมะเร็งเม็ดเลือดขาว พบว่าการลดมิติข้อมูลด้วย Linear SVM Weight ให้ประสิทธิภาพดี

ที่สุด นอกจากนี้ยังได้มีการประยุกต์ใช้กิริตอีอัลกอริทึม เพื่อคัดเลือกคุณลักษณะที่สำคัญของข้อมูล

นิเวศ [9] ได้เสนอแบบจำลองการจัดหมวดหมู่เอกสาร ภาษาไทยโดยลดคุณลักษณะของเอกสารภาษาไทย ด้วยวิธี Information Gain วัดประสิทธิภาพพบว่าอัลกอริทึม SVM ให้ประสิทธิภาพสูงสุด เมื่อพิจารณาการคัดเลือกคุณลักษณะจากกลุ่มตัวอย่างของ SVM พบว่าสามารถลดขนาดลงได้มากถึง 90 %

3. วิธีการดำเนินการวิจัย

กระบวนการเปรียบเทียบวิธีการเลือกคุณลักษณะที่เหมาะสม เพื่อการจัดหมวดหมู่เว็บเพจผิดกฎหมาย แสดงไว้ในภาพที่ 1



ภาพที่ 1: แบบจำลองการเปรียบเทียบวิธีการเลือกคุณลักษณะที่เหมาะสม เพื่อการจัดหมวดหมู่เว็บเพจผิดกฎหมาย

3.1 การเตรียมข้อมูล

ในการวิจัยใช้ข้อมูลเว็บเพจจากฐานข้อมูลเว็บเพจ[11] ที่มีการเก็บไว้โดยมีจำนวนเว็บเพจ 5,399 URL โดยการ

แบ่งประเภทข้อมูลออกเป็น 4 ประเภท คือ อนาจาร (Indecent) การพนัน (Gamble) ไม่เหมาะสม (Inappropriate) และทั่วไป (Normal)

3.2 การแทนเว็บเพจ (Web Page Representation)

การแทนค่าเว็บเพจและข้อมูลสืบค้นหรือข้อมูลนำเข้า (Query) ด้วยเวกเตอร์ที่เรียกว่า เวกเตอร์ของคำ (Vector of Terms) โดยแกนแต่ละแกนของเวกเตอร์ คือ คำ (Terms) ที่แสดงอยู่ในเอกสารคำเหล่านี้ได้จากการนำคำทั้งหมดที่ปรากฏอยู่ในทุกเอกสารเพื่อใช้อ้างอิงในการสร้างเวกเตอร์ของคำ ที่ตรงกัน และความยาวของเวกเตอร์ คือ ค่าความถี่ของคำที่ถูกเลือก ซึ่งปรากฏอยู่ในเอกสารนั้นประกอบรวมกัน

3.3 การกำหนดชุดข้อมูล (Defining Dataset)

การกำหนดจำนวนข้อมูลเว็บเพจที่ใช้เป็นชุดฝึกฝนสำหรับสร้างโมเดล และชุดทดสอบสำหรับทดสอบโมเดลที่สร้างขึ้น โดยระบบจะสุ่มข้อมูลเว็บเพจตามจำนวนเว็บเพจที่กำหนดไว้ จากนั้นจะแบ่งชุดข้อมูลออกเป็น 10 ชุดย่อย ซึ่งแต่ละชุดย่อยมีข้อมูลเว็บเพจหมวดหมู่ละเท่า ๆ กัน เพื่อเตรียมไว้ในขั้นตอนการประเมินผลโมเดลโดยวิธี 10-Fold Cross Validation

3.4 การคัดเลือกคุณลักษณะ (Feature Selection)

เป็นการเลือกข้อมูลที่มีประโยชน์ต่อการจัดหมวดหมู่เว็บเพจและการลดขนาดของมิติเอกสารลง โดยในการวิจัยนี้ใช้วิธีการจัดอันดับแบบ information gain, gain ratio และ chi-square โดยกระบวนการนี้ช่วยลดค่าที่ไม่มีความสำคัญ และช่วยให้การประมวลผลมีความเร็วมากขึ้น

3.5 การทดสอบอัลกอริทึมในการจัดหมวดหมู่ (Classification Algorithm)

การทดสอบอัลกอริทึมในการจัดหมวดหมู่ โดยนำข้อมูลชุดฝึกฝนที่ผ่านการคัดเลือกคุณลักษณะแล้วมาเข้ากระบวนการจัดหมวดหมู่ โดยเทคนิคซัพพอร์ตเวกเตอร์แมชชีน เพื่อวัดประสิทธิภาพและเปรียบเทียบประสิทธิภาพของเทคนิคการคัดเลือกคุณลักษณะ

3.6 Classification Model

เมื่อทำการทดสอบอัลกอริทึมที่ใช้ในการจัดหมวดหมู่เว็บเพจหลังจากการคัดเลือกคุณลักษณะในข้างต้น จะได้โมเดลในการจัดหมวดหมู่เว็บเพจผิดกฎหมาย ที่มีความเหมาะสมที่สุด

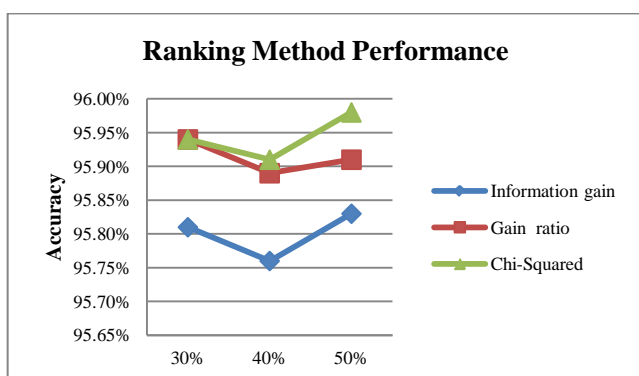
4. ผลการดำเนินงาน

4.1. ผลการเปรียบเทียบประสิทธิภาพการคัดเลือกคุณลักษณะ

โดยวิธีการจัดอันดับแบบต่าง ๆ และการจัดหมวดหมู่ด้วย SVM ในส่วนนี้เป็นการทดสอบความถูกต้องของการจัดหมวดหมู่ด้วยเทคนิคซัพพอร์ตเวกเตอร์แมชชีน โดยการใช้ข้อมูลที่ผ่านการคัดเลือกคุณลักษณะด้วยวิธี Information gain, Gain ratio และ Chi-Square เพื่อค้นหาวิธีการที่ดีที่สุดได้รายละเอียดดังตารางที่ 1 โดยวิธีการจัดอันดับแบบ Chi-Square ให้ผลลัพธ์ที่ดีที่สุดคือ เมื่อลดจำนวนข้อมูล 50% จากข้อมูลทั้งหมด 160 แอตทริบิวต์ ได้ค่าความถูกต้อง (Accuracy) ที่ร้อยละ 95.98 รองลงมาคือวิธีการจัดอันดับแบบ Gain ratio โดยลดจำนวนข้อมูล 30% คือ 112 แอตทริบิวต์ และวิธีการจัดอันดับแบบ Information gain โดยลดจำนวนข้อมูล 50% คือ 80 แอตทริบิวต์ ได้ค่าความถูกต้อง (Accuracy) ที่ร้อยละ 95.94 และ 95.83 ตามลำดับ

ตารางที่ 1: การเปรียบเทียบประสิทธิภาพของเทคนิคการคัดเลือกคุณลักษณะด้วยวิธีการจัดอันดับแบบต่าง ๆ

Ranking Method	จำนวนการลดแอตทริบิวต์		
	30%	40%	50%
Information gain	95.81%	95.76%	95.83%
Gain ratio	95.94%	95.89%	95.91%
Chi-Square	95.94%	95.91%	95.98%



ภาพที่ 2: กราฟแสดงผลลัพธ์วิธีการคัดเลือกคุณลักษณะด้วยวิธีการจัดอันดับแบบต่าง ๆ

จากภาพที่ 1 กราฟแสดงให้เห็นว่าผลลัพธ์ของการคัดเลือกคุณลักษณะด้วยวิธีการจัดอันดับแบบ Chi-Squared ให้ประสิทธิภาพดีที่สุด

4.2. ผลการทดสอบความถูกต้องของการจัดหมวดหมู่

การคัดเลือกคุณลักษณะด้วยวิธีการจัดอันดับแบบต่าง ๆ ด้วยอัลกอริทึม SVM kernel แบบ RBF ในส่วนนี้เป็นการนำโครงสร้างของ SVM มาทดสอบกับข้อมูลที่ผ่านการคัดเลือกคุณลักษณะ โดยจะทำการปรับเปลี่ยนค่า gamma ที่เหมาะสมสำหรับ kernel แบบ RBF เพื่อทดสอบหาค่าที่เหมาะสมที่ให้ประสิทธิภาพสูงสุด รายละเอียดดังตารางที่ 2, 3 และ 4

ตารางที่ 2: ผลการทดสอบประสิทธิภาพของเทคนิค Information gain โดยการหาค่าที่เหมาะสมสำหรับ kernel แบบ RBF

Gamma	จำนวนการลดแอตทริบิวต์		
	30%	40%	50%
0.8	97.76	97.72	97.57
0.9	97.76	97.72	97.70
1.0	97.76	97.80	97.67
1.1	97.83	97.80	97.61
1.2	97.80	97.81	97.61

ตารางที่ 3: ผลการทดสอบประสิทธิภาพของเทคนิค Gain ratio โดยการหาค่าที่เหมาะสมสำหรับ kernel แบบ RBF

Gamma	จำนวนการลดแอตทริบิวต์		
	30%	40%	50%
0.8	97.89	97.91	97.67
0.9	97.89	97.91	97.70
1.0	97.89	97.91	97.68
1.1	97.91	97.89	97.74
1.2	97.89	97.89	97.72

ตารางที่ 4: ผลการทดสอบประสิทธิภาพของเทคนิค Chi-Square โดยการหาค่าที่เหมาะสมสำหรับ kernel แบบ RBF

Gamma	จำนวนการลดแอตทริบิวต์		
	30%	40%	50%
0.8	97.89	97.85	97.78
0.9	97.89	97.85	97.80
1.0	97.89	97.89	97.74
1.1	97.93	97.87	97.67
1.2	97.89	97.87	97.67

5. สรุปผลและข้อเสนอแนะ

การเปรียบเทียบประสิทธิภาพของการคัดเลือกคุณลักษณะที่เหมาะสมในการจัดหมวดหมู่เว็บเพจผิดกฎหมาย เพื่อค้นหาวิธีการที่เหมาะสมในการจัดหมวดหมู่เว็บเพจผิดกฎหมายโดยใช้เทคนิคซัพพอร์ตเวกเตอร์แมชชีนเคอร์เนลฟังก์ชันแบบเรเดียลเบสิสฟังก์ชัน สรุปได้ว่า การคัดเลือกคุณลักษณะสามารถลดจำนวนของแอตทริบิวต์ได้และสามารถเพิ่มประสิทธิภาพของการจัดหมวดหมู่เว็บเพจ จากการเปรียบเทียบประสิทธิภาพการคัดเลือกคุณลักษณะแบบต่างๆ พบว่าวิธีการจัดอันดับแบบ Chi-Square ให้ผลลัพธ์ที่ดีที่สุด คือ เมื่อลดจำนวนแอตทริบิวต์ลง 50% ได้ค่าความถูกต้อง (Accuracy) ที่ร้อยละ 95.98 และเมื่อทำการปรับค่าพารามิเตอร์ค่าพารามิเตอร์ Gamma เป็น 0.8,0.9,1.0,1.1 และ 1.2 ผลการทดลองพบว่าค่าพารามิเตอร์ Gamma ส่งผลให้ค่าความถูกต้องเพิ่มขึ้นโดยวิธีการคัดเลือกคุณลักษณะด้วยวิธีการจัดอันดับแบบ Chi-Square ให้ค่าความถูกต้องเพิ่มขึ้นจากร้อยละ 95.94 เป็น 97.93 ที่การลดจำนวน แอตทริบิวต์ 30% และค่าพารามิเตอร์ Gamma ที่เหมาะสม เท่ากับ 1.1

การวิจัยในการคัดเลือกคุณลักษณะข้อมูลในอนาคต อาจทำการเปรียบเทียบเทคนิคการคัดเลือกคุณลักษณะด้วยวิธีการจัดอันดับและวิธีการค้นหาแบบไปข้างหน้า เพื่อหาเทคนิคที่ดีที่สุดต่อไป

เอกสารอ้างอิง

- [1] วิภาวรรณ บัวทอง, “การเปรียบเทียบประสิทธิภาพของเทคนิคการลดมิติข้อมูลด้วยวิธีการจัดอันดับแบบ Information Gain, Gain Ratio และ Linear SVM Weights” *การประชุมวิชาการทางเทคโนโลยีคอมพิวเตอร์และระบบสารสนเทศประยุกต์ ครั้งที่ 5*. 185-189, 2555.
- [2] จิราภรณ์ ถมแก้ว และ ศรีณย์ อินทโกสุม, “การจำแนกข้อมูลโดยการคัดเลือกคุณลักษณะที่สำคัญ” *การประชุมวิชาการเสนอผลงานวิจัยระดับบัณฑิตศึกษาแห่งชาติ ครั้งที่ 23*. หน้า 7-12 . มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน, 2554.
- [3] วาทีน น้อยเพียร และ พยุง มีสัง, “การเปรียบเทียบเทคนิคการคัดเลือกคุณลักษณะแบบการกรองและการควมรวมของการทำเหมืองข้อความเพื่อการจำแนกข้อความ,” *วารสารวิชาการเทคโนโลยีอุตสาหกรรมปีที่ 9 ฉบับที่ 3*, 2556.
- [4] ภรณ์ชา อามฤรัตน์ และ พยุง มีสัง, “การเปรียบเทียบประสิทธิภาพการจำแนกกลุ่มข้อมูลของโรคลมร้อนด้วยวิธีการทางเครือข่ายประสาทเทียม” *การประชุมทางวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศครั้งที่ 6*, 2553.
- [5] ชาญพร ศรีดอกไม้, “การเปรียบเทียบประสิทธิภาพการจำแนกเว็บเพจระหว่างระหว่างเทคนิค ซัพพอร์ตเวกเตอร์แมชชีน นาอิมพ์ คัดสันใจและ กฎพื้นฐาน ร่วมกับวิธีการกรอง” *การประชุมทางวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศครั้งที่ 6*, 2553.
- [6] สุกเทพ สติมัน และ ทิพยา จินตโกวิท, “การวัดประสิทธิภาพของอัลกอริทึมในการจำแนกประเภทข้อความสั้นกรณีศึกษา ข้อความภาษาไทยบนเว็บไซต์ทวิตเตอร์” *การประชุมทางวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศครั้งที่ 7*, 2554.
- [7] นิยม กุทธิ, “การจัดหมวดหมู่เว็บเพจภาษาไทยโดยใช้คุณลักษณะที่ปรากฏบนหน้าเว็บเพจ.” สารนิพนธ์ระดับบัณฑิตศึกษาสาขาวิชาระบบสารสนเทศเพื่อการจัดการ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2553.
- [8] จิรชา ไทยสังคม, “การพัฒนาแบบจัดหมวดหมู่เอกสารบนระบบเครือข่ายโดยใช้โครงข่ายประสาทเทียม.” สารนิพนธ์วิทยาศาสตร์มหาบัณฑิต สาขาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2549.
- [9] นิเวศ จิระวิชิตชัย ปริญญา สงวนศักดิ์ และ พยุง มีสัง, “การจัดหมวดหมู่เอกสารภาษาไทยแบบอัตโนมัติด้วยซัพพอร์ตเวกเตอร์แมชชีน,” *การประชุมทางวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศครั้งที่ 6*, 2553.
- [10] วงกต ศรีอุไร พยุง มีสัง และ ชูชาติ หฤไชยะศักดิ์, “การเตรียมฟิเจอร์บนพื้นฐานแบบจำลองหัวข้อสำหรับการจำแนกหมวดหมู่ของเอกสาร” *การประชุมทางวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 5*, 2552.
- [11] ก่อเกียรติ แก้วกิ่ง และ กาญจนา วิริยะพันธ์, “การจำแนกเว็บเพจไม่เหมาะสมและการวิเคราะห์ความเห็นบนเว็บโดยใช้โครงข่ายประสาทเทียม” *การประชุมทางวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 9*, 2556.

ผลการสังเคราะห์รูปแบบการพัฒนาทักษะการออกแบบขั้นตอนวิธีทางคอมพิวเตอร์ โดยใช้เทคโนโลยีเว็บ 2.0 ด้วยเทคนิคพี่เลี้ยง

The Synthesis of Skill Development Model of Computing Design Process by Using 2.0 Web Technology and Mentoring Technique

สุนันทา กลิ่นถาวร (Sununta Klintawon)¹ เนตรชนก จันทร์สว่าง(Netchanok Jansawang)²

วรปภา อารีราษฎร์ (Worapapha Arreerard)³

¹นักศึกษาคณะศึกษาศาสตร์ สาขาวิชาคอมพิวเตอร์ศึกษา มหาวิทยาลัยราชภัฏมหาสารคาม

²คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏมหาสารคาม

³คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏมหาสารคาม

Sununta101@hotmail.com, jansawangn@yahoo.com, dr.worapapha@hotmail.com

บทคัดย่อ

การวิจัยครั้งนี้เป็นการวิจัยเพื่อสังเคราะห์รูปแบบการพัฒนาทักษะการออกแบบขั้นตอนวิธีทางคอมพิวเตอร์ โดยใช้เทคโนโลยีเว็บ 2.0 ด้วยเทคนิคพี่เลี้ยงมีวัตถุประสงค์เพื่อ สังเคราะห์รูปแบบและกิจกรรมการพัฒนาทักษะการออกแบบขั้นตอนวิธีทางคอมพิวเตอร์ โดยใช้เทคโนโลยีเว็บ 2.0 ด้วยเทคนิคพี่เลี้ยง กลุ่มเป้าหมายที่ใช้ในการวิจัยครั้งนี้ เป็นผู้เชี่ยวชาญ 3 กลุ่ม คือ 1)กลุ่มการจัดการเรียนการสอนในสาขาคอมพิวเตอร์ 2)กลุ่มนักวิชาการทั่วไปและ 3)กลุ่มเทคโนโลยีสารสนเทศ รวมทั้งสิ้นจำนวน 12 คน โดยใช้การประสมกลุ่มย่อยเครื่องมือที่ใช้ในการวิจัยได้แก่ แบบสอบถามสถิติที่ใช้ได้แก่ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน

ผลการวิจัยพบว่า 1) ผลการสังเคราะห์รูปแบบการพัฒนาทักษะการออกแบบขั้นตอนวิธีทางคอมพิวเตอร์ มีองค์ประกอบ 4 ส่วน คือ 1.1) หลักการที่เกี่ยวข้อง 1.2)แนวคิด ทฤษฎี 1.3) การพัฒนาทักษะ และ 1.4) เทคโนโลยีสารสนเทศและการสื่อสาร โดยความคิดเห็นของผู้เชี่ยวชาญที่มีต่อองค์ประกอบของรูปแบบโดยรวมอยู่ในระดับมากที่สุด และ 2) ผลการสังเคราะห์รูปแบบกิจกรรมการพัฒนาทักษะการออกแบบขั้นตอนวิธีทางคอมพิวเตอร์ โดยใช้เทคโนโลยีเว็บ 2.0 ด้วยเทคนิคพี่เลี้ยง มีกิจกรรมการดำเนินงาน 3 ระยะ คือ ระยะที่ 1 การเตรียมความพร้อม ระยะที่ 2 การพัฒนาทักษะ และ ระยะที่ 3 การติดตามผล โดยความคิดเห็นของผู้เชี่ยวชาญที่มีต่อกิจกรรมโดยรวมอยู่ในระดับมากที่สุด

คำสำคัญ การพัฒนาทักษะ การออกแบบขั้นตอนวิธีทางคอมพิวเตอร์ เทคโนโลยีเว็บ 2.0 เทคนิคพี่เลี้ยง

Abstract

The research aimed at synthesizing a skill development model and learning activities for computing design process by using 2.0 web technology and a mentoring technique. The target expert population was twelve participants consisting of experts in computer education, information technology and academicians. The instrument was a questionnaire. The statistics used were mean and standard deviation.

The research showed that a skill development model for computing design process by using 2.0 web technology and a mentoring technique consisted of four major components: 1) rationale, 2) concept and theory, 3) skill development and 4) information technology and communication. The finding showed that the average level of the opinions of experts towards the skill development model was very high. In addition, the research showed that the model of computing design process by using 2.0 web technology and a mentoring technique was divided into three steps: 1) preparation, 2) skill development and 3) evaluation. The finding indicated that the average level of the opinion of the experts towards the training activities of the model was very high

Keywords: Skill Development, Computing Process

Design, 2.0 Web Technology, Mentoring Technique

1. บทนำ

ขั้นตอนวิธีหรือที่เรียกว่า อัลกอริทึม (Algorithm) เป็นลำดับของขั้นตอนเพื่อใช้แก้ปัญหาและเมื่อปฏิบัติตามขั้นตอนอย่างเป็นลำดับ สามารถให้ผลลัพธ์ตามที่ต้องการ [1]ซึ่งถือได้ว่าขั้นตอนวิธีเป็นวิธีการคิดอย่างเป็นระบบเพื่อใช้ในการแก้ปัญหาโจทย์ทางคอมพิวเตอร์ในรูปแบบหนึ่งดังนั้นขั้นตอนวิธีจึงเป็นพื้นฐานที่จำเป็นอย่างยิ่งสำหรับการเขียนโปรแกรมคอมพิวเตอร์

จากการสำรวจสภาพปัญหาของนักเรียนที่จะเข้าเรียนในระดับปริญญาตรี คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏมหาสารคาม พบว่า นักเรียนขาดความรู้พื้นฐาน และทักษะทางการเขียนโปรแกรมและการออกแบบขั้นตอนวิธีทางคอมพิวเตอร์ ทั้งนี้จากการสำรวจได้เสนอแนะแนวทางในการเตรียมความพร้อมให้กับผู้เรียนก่อนเข้าศึกษาหรือระหว่างการศึกษานานชั้นปีที่ 1 เพื่อเป็นการเตรียมความพร้อมให้กับนักศึกษา ซึ่งตรงกับข้อมูลความคิดเห็นของอาจารย์ผู้สอนรายวิชาการเขียนโปรแกรมภาษาคอมพิวเตอร์ 1 จากมหาวิทยาลัยราชภัฏ 22 แห่ง ที่พบว่านักศึกษามีความรู้พื้นฐานในการเขียน โปรแกรมและอัลกอริทึมน้อย ดังนั้นจึงได้เสนอให้มีการส่งเสริมให้นักศึกษาได้เรียนรู้ทางด้านอัลกอริทึม เพื่อให้นักศึกษาได้เข้าใจหลักการเขียนโปรแกรมอย่างชัดเจน [2]

Scratch เป็น โปรแกรมภาษาคอมพิวเตอร์ภาษาหนึ่งในเทคโนโลยีเว็บ 2.0 ที่ส่งเสริมให้เด็กใช้ความคิดวิเคราะห์คิดอย่างเป็นระบบคิดเชิงตรรกะและคิดสร้างสรรค์เพื่อแก้ปัญหาทางคอมพิวเตอร์ด้วยการสร้างโปรแกรมผลงาน[3] โดยกระบวนการทำงานของ Scratchเหมาะสมกับการเรียนรู้โดยใช้ปัญหาเป็นฐาน (Problem Based Learning : PBL)ซึ่งเป็นวิธีการเรียนการเรียนรู้ที่เน้นให้ผู้เรียนได้เกิดการเรียนรู้ด้วยตนเอง โดยใช้ปัญหาเป็นเครื่องกระตุ้น การจัดการเรียนการสอนโดยใช้ปัญหาเป็นฐานไม่เพียงแต่จะเพิ่มความรู้ ยังมีกระตุ้นต่อสื่อสารและมีกลไกการทำงานกลุ่ม[4] ซึ่งเหมาะสมกับการนำมาเป็นรูปแบบเพิ่มทักษะให้กับผู้เรียนในระดับอุดมศึกษา เพราะผู้เรียนในระดับนี้ชอบการค้นคว้า การหาคำตอบ และยังเป็นวัยที่มีความต้องการในการที่จะเป็นที่ยอมรับของเพื่อนและสังคมรอบ ๆ ตัว

จากการศึกษาและสังเคราะห์เอกสาร และงานวิจัยที่เกี่ยวข้อง จึงได้สังเคราะห์รูปแบบการพัฒนาทักษะการออกแบบขั้นตอนวิธีทางคอมพิวเตอร์ โดยใช้เทคโนโลยีเว็บ 2.0ด้วยเทคนิคที่เรียบง่ายเพื่อ

เป็นกระบวนการในการเตรียมความพร้อมเพื่อให้นักศึกษามีทักษะในการออกแบบขั้นตอน ซึ่งเป็นความรู้พื้นฐานในการเรียนในหลักสูตรสาขาคอมพิวเตอร์ได้อย่างมีประสิทธิภาพ

2. วรรณกรรมที่เกี่ยวข้อง

ตารางที่ 1 : การสังเคราะห์วิธีการออกแบบขั้นตอนวิธีทางคอมพิวเตอร์

ทฤษฎี \ กระบวนการ	นิยามปัญหา	เขียนอัลกอริทึม	เขียนผังงาน	เขียนโปรแกรม
พิสุทธา อารีราษฎร์	✓	✓	✓	
อุไรวรรณ โสกา	✓	✓		
ประกาศกร เขียงทอง	✓	✓	✓	
จิรพร วีระพันธุ์	✓	✓	✓	✓
ธีรภัทร เกษสกุล	✓	✓	✓	✓

2.1 การออกแบบขั้นตอนวิธีทางคอมพิวเตอร์เป็นกระบวนการวิธีการกำหนดลัดลำดับของวิธีการซึ่งให้ผลลัพธ์สำหรับปัญหาต่าง ๆ ในรูปของขั้นตอนที่มีจำนวนจำกัดขั้นตอนวิธีได้นำมาใช้ในการดำเนินการเขียน โปรแกรมคอมพิวเตอร์ เพื่อวิเคราะห์แยกแยะวิธีการทำงานให้เป็น [5] ซึ่งกระบวนการในการเขียนขั้นตอนวิธี เป็นการเขียนขั้นตอนเพื่อสั่งคอมพิวเตอร์ทำงานให้ตรงตามต้องการ ภายใต้อัจฉริยะของคอมพิวเตอร์

2.2 การจัดการเรียนรู้โดยใช้ปัญหาเป็นฐานเป็นการสอนแบบใช้ปัญหาเป็นฐาน โดยปัญหาจะเป็นตัวกระตุ้นให้ผู้เรียนต้องไปแสวงหาความรู้ความเข้าใจด้วยตนเอง[7] โดยผู้เรียนจะสามารถสร้างความรู้ใหม่จากการใช้ปัญหาที่เกิดขึ้นเป็นบริบท (Context) ของการเรียนรู้เพื่อให้ผู้เรียนเกิดทักษะในการคิดวิเคราะห์และคิดแก้ปัญหา

ตารางที่ 2 : การสังเคราะห์การจัดการเรียนรู้โดยใช้ปัญหาเป็นฐาน

ทฤษฎี \ กระบวนการ	ปัญหา	การบูรณาการ	การทำงานเป็น	การแก้ปัญหา	การเรียนรู้ด้วยตนเอง
บาร์ โวส์และทัมบลิ้น	✓	✓		✓	✓
กูค	✓	✓		✓	✓
สมบัติ เฝ้าพงศ์คล้าย	✓	✓			✓
กมลวรรณ สายจุฑาวัดน์	✓	✓	✓	✓	✓
อานุกาภ เลขะกุล	✓	✓			✓
มหาวิทยาลัยอีสเทิร์นเอเชีย	✓	✓			
สมทรง สิทธิ	✓	✓			✓

2.3 เทคโนโลยีเว็บ 2.0 คือระบบที่ผู้ใช้เป็นผู้สร้างและแก้ไขเนื้อหาไม่ว่าจะเป็นเนื้อหาที่สร้างคนเดียวและเผยแพร่ให้คนอื่นได้อ่านหรือเนื้อหาสาธารณะที่ทุกคนสามารถมีส่วนร่วมช่วยกันสร้างจึงทำให้ปริมาณและความหลากหลายของเนื้อหาในสังคมมีมากขึ้น [3] เพราะทุกคนสามารถเป็นผู้สร้างเนื้อหาได้และผู้ใช้เป็นผู้กำหนดคุณค่าของเนื้อหาไม่ว่าจะด้วยการให้คะแนนการโหวตหรือแม้แต่การค้นหาข้ามเว็บไซต์ผ่าน Search Engine

2.4 เทคนิคพี่เลี้ยงถูกนำมาใช้เป็นเครื่องมือของการจัดการความรู้ในแง่ของการถ่ายทอดความรู้ระหว่างผู้ที่ประสบการณ์มากกว่าและผู้ที่มีประสบการณ์น้อยกว่า [6] ระบบพี่เลี้ยงเป็นรูปแบบหนึ่งของการแลกเปลี่ยนความรู้ที่มีจุดประสงค์พื้นฐานในการสร้างปฏิสัมพันธ์อย่างไม่เป็นทางการระหว่างพี่เลี้ยง (Mentor) กับผู้ที่ถูกสอนงาน (Mentee)

ระบบพี่เลี้ยงได้รับการพิสูจน์ว่าเป็นระบบที่มีประสิทธิภาพสูงและนิยมใช้กันมากซึ่งจัดได้ว่าเป็นช่องทางที่ดีสำหรับการพัฒนาศักยภาพของบุคลากรในองค์กร

3. วิธีการดำเนินงาน

3.1 กลุ่มเป้าหมาย

กลุ่มเป้าหมายที่ใช้ในการวิจัยครั้งนี้ เป็นผู้เชี่ยวชาญ 3 กลุ่ม คือ 1) กลุ่มการจัดการเรียนการสอนในสาขาคอมพิวเตอร์ 2) กลุ่มนักวิชาการทั่วไป และ 3) กลุ่มเทคโนโลยีสารสนเทศ กลุ่มละ 4 คน รวมทั้งสิ้นจำนวน 12 คน

3.2 เครื่องมือที่ใช้ในการวิจัย

แบบสอบถามผู้เชี่ยวชาญที่มีต่อรูปแบบการพัฒนาทักษะการออกแบบขั้นตอนวิธีทางคอมพิวเตอร์ โดยใช้เทคโนโลยีเว็บ 2.0 ด้วยเทคนิคพี่เลี้ยง

3.3 ขั้นตอนการดำเนินงาน

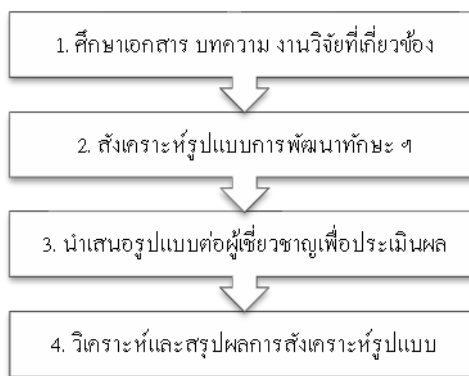
3.3.1 การศึกษาเอกสาร บทความ และงานวิจัยที่เกี่ยวข้อง ผู้วิจัยได้ทำการวิเคราะห์ข้อมูลและจัดหมวดหมู่ของความสัมพันธพบว่า ในการพัฒนารูปแบบการพัฒนาทักษะการออกแบบขั้นตอนวิธีทางคอมพิวเตอร์ โดยใช้เทคโนโลยีเว็บ 2.0 ด้วยเทคนิคพี่เลี้ยงนั้น ควรจะมีองค์ประกอบทั้งหมด 4 ส่วน คือ 1) หลักการที่เกี่ยวข้อง 2) แนวคิด ทฤษฎี 3) การพัฒนาทักษะ และ 4) เทคโนโลยีเว็บ 2.0

3.3.2 สังเคราะห์รูปแบบการพัฒนาทักษะฯ ผู้วิจัยได้ทำการร่างรูปแบบฯ และกำหนดรายละเอียดร่างรูปแบบ โดยแบ่งเป็น 2 ส่วน

คือ 1) รูปแบบการพัฒนาทักษะการออกแบบขั้นตอนวิธีทางคอมพิวเตอร์ โดยใช้เทคโนโลยีเว็บ 2.0 ด้วยเทคนิคพี่เลี้ยง และ 2) รูปแบบกิจกรรมการพัฒนาทักษะการออกแบบขั้นตอนวิธีทางคอมพิวเตอร์ โดยใช้เทคโนโลยีเว็บ 2.0 ด้วยเทคนิคพี่เลี้ยง

3.3.3 นำเสนอรูปแบบต่อผู้เชี่ยวชาญเพื่อประเมินผลผู้วิจัยนำร่างรูปแบบที่พัฒนาขึ้นเสนอต่อผู้เชี่ยวชาญ เพื่อขอคำแนะนำและประเมินรูปแบบที่พัฒนาขึ้น

3.3.4 วิเคราะห์และสรุปผลการสังเคราะห์รูปแบบผู้วิจัยได้นำผลการประเมินโดยผู้เชี่ยวชาญมากทำการวิเคราะห์ และสรุปผลตามกระบวนการของการวิจัย

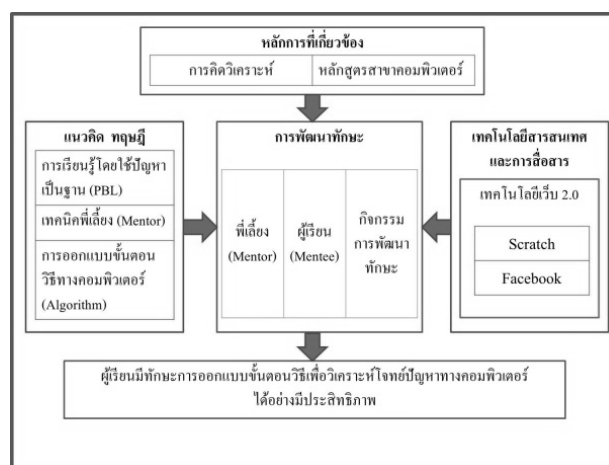


ภาพที่ 1: ขั้นตอนการดำเนินงานสังเคราะห์รูปแบบฯ

3.4 การสังเคราะห์ข้อมูล

การวิเคราะห์ข้อมูล ผู้วิจัยนำแบบสอบถามความคิดเห็นของผู้เชี่ยวชาญที่ได้จากการประเมินรูปแบบ มาวิเคราะห์ความคิดเห็น โดยใช้สถิติค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน

4. ผลการดำเนินงาน



ภาพที่ 2: รูปแบบการพัฒนาทักษะการออกแบบขั้นตอนวิธีทางคอมพิวเตอร์ โดยใช้เทคโนโลยีเว็บ 2.0 ด้วยเทคนิคพี่เลี้ยง

4.1 ผลการสังเคราะห์รูปแบบการพัฒนาทักษะการออกแบบ ขั้นตอนวิธีทางคอมพิวเตอร์ โดยใช้เทคโนโลยีเว็บ 2.0 ด้วย เทคนิคที่เลี้ยง ผลการดำเนินงาน พบว่า

4.1.1 นิยาม รูปแบบการพัฒนาทักษะ คือ องค์ประกอบ การพัฒนาทักษะ เพื่อเป็นการเตรียมความพร้อมให้กับนักศึกษา ในการออกแบบขั้นตอนวิธี และนำไปเป็นแนวทางในการแก้ โจทย์ปัญหาคอมพิวเตอร์ ที่มุ่งเน้นให้ผู้เรียนมีทักษะการคิดอย่างเป็นระบบ

4.1.2 องค์ประกอบของรูปแบบ มีองค์ประกอบ 4 ส่วน ดัง ภาพที่ 2

1) หลักการที่เกี่ยวข้อง ประกอบด้วย แนวคิดของการ คิดวิเคราะห์ ซึ่งหมายถึง ความสามารถในการพิจารณาแยก ส่วนย่อยๆ และหาความสัมพันธ์ของแต่ละองค์ประกอบว่า สัมพันธ์กันอย่างไร และหลักสูตรสาขาคอมพิวเตอร์ ที่บ่งบอก ถึงขอบเขตองค์ความรู้ของสาขาคอมพิวเตอร์

2) แนวคิด ทฤษฎี เป็นการนำทฤษฎีการเรียนรู้ต่างๆ มาเป็นแนวทางจัดการเรียนรู้ซึ่งประกอบด้วย การจัดการเรียนรู้ โดยใช้ปัญหาเป็นฐาน การเรียนรู้ร่วมกันโดยใช้เทคนิคที่เลี้ยง และแนวคิดการออกแบบขั้นตอนวิธีทางคอมพิวเตอร์

3) การพัฒนาทักษะ เป็นส่วนที่ระบุถึงปัจจัยของการ พัฒนาทักษะไว้ 3 อย่างคือ 1) ผู้เรียนหรือผู้เข้าร่วมการพัฒนา ทักษะ 2) พี่เลี้ยงซึ่งจะคอยทำหน้าที่ให้คำแนะนำและการ ช่วยเหลือผู้เรียน และ 3) กิจกรรมการพัฒนาทักษะการออกแบบ ขั้นตอนวิธีทางคอมพิวเตอร์ โดยใช้เทคโนโลยีเว็บ 2.0 ด้วย เทคนิคที่เลี้ยง

4) เทคโนโลยีสารสนเทศและการสื่อสารเป็นการนำ เทคโนโลยีเข้ามาใช้เป็นเครื่องมือใน 2 ลักษณะคือ 1) เป็น เครื่องมือในการพัฒนาทักษะ และ 2) เป็นเครื่องมือในการ ติดต่อสื่อสาร

4.1.3 ผลการศึกษาความคิดเห็นของผู้เชี่ยวชาญที่มีต่อ รูปแบบการพัฒนาทักษะการออกแบบขั้นตอนวิธี ทาง คอมพิวเตอร์ โดยใช้เทคโนโลยีเว็บ 2.0 ด้วยเทคนิคที่เลี้ยงจากการ ประชุมกลุ่มย่อย และการประเมินรูปแบบ โดยผู้เชี่ยวชาญ ดังแสดง ในตารางที่ 3

ตารางที่ 3: ความคิดเห็นของผู้เชี่ยวชาญที่มีต่อรูปแบบการพัฒนา ทักษะ

รายการ	ระดับความคิดเห็น		
	\bar{X}	S.D.	ความหมาย
ความเหมาะสมของ องค์ประกอบของรูปแบบ	4.75	0.45	มากที่สุด
1.1 หลักการที่เกี่ยวข้อง	4.71	0.55	มากที่สุด
1.2 แนวคิด ทฤษฎี	4.83	0.38	มากที่สุด
1.3 การพัฒนาทักษะ	4.53	0.51	มากที่สุด
1.4 เทคโนโลยีสารสนเทศ และการสื่อสาร	4.83	0.38	มากที่สุด

จากตารางที่ 3 พบว่า ความคิดเห็นของผู้เชี่ยวชาญที่มีต่อความ เหมาะสมขององค์ประกอบของรูปแบบอยู่ในระดับมากที่สุด ($\bar{X}=4.75$, $S.D.=0.45$) และเมื่อพิจารณาเป็นรายด้านพบว่า ด้าน เทคโนโลยีสารสนเทศและการสื่อสาร และด้านแนวคิด ทฤษฎี มี ระดับความเหมาะสมเท่ากัน คือมีความเหมาะสมมากที่สุด ($\bar{X}=4.83$, $S.D.=0.38$) รองลงมาคือด้าน หลักการที่เกี่ยวข้อง ($\bar{X}=4.71$, $S.D.=0.55$) และด้านการพัฒนาทักษะ ($\bar{X}=4.53$, $S.D.=0.51$) ตามลำดับ

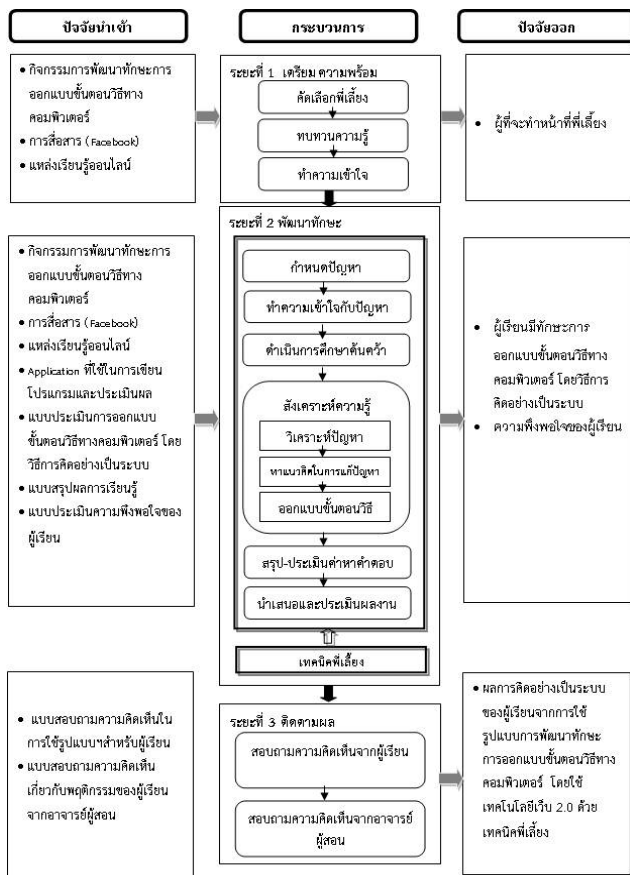
4.2 ผลการสังเคราะห์รูปแบบกิจกรรมการพัฒนาทักษะการ ออกแบบขั้นตอนวิธีทางคอมพิวเตอร์ โดยใช้เทคโนโลยีเว็บ 2.0 ด้วยเทคนิคที่เลี้ยง ผลการดำเนินงาน พบว่า

4.2.1 นิยาม รูปแบบกิจกรรม หมายถึง แบบแผนที่ แสดงถึงความสัมพันธ์ของกระบวนการหรือขั้นตอนการฝึกอบรม การวิเคราะห์และออกแบบขั้นตอนวิธีทางคอมพิวเตอร์

4.2.2 กิจกรรมการดำเนินงานของรูปแบบกิจกรรม มี 3 ระยะ ดังภาพที่ 3

1) ระยะที่ 1 เตรียมความพร้อม ประกอบด้วย 3 ขั้นตอน คือ 1. คัดเลือกพี่เลี้ยง 2. ทบทวนความรู้ และ 3. ทำความเข้าใจ

2) ระยะที่ 2 พัฒนาทักษะ ในระยะนี้จะนำเทคนิคที่ เลี้ยงเข้ามามีส่วนร่วมในการจัดการพัฒนาทักษะ โดยพี่เลี้ยงจะทำ หน้าที่ตามบทบาทหน้าที่ที่ได้ระบุไว้ ซึ่งจะประกอบด้วย 6 ขั้นตอนคือ 1. กำหนดปัญหา 2. ทำความเข้าใจกับปัญหา 3. ดำเนินการศึกษาค้นคว้า 4. สังเคราะห์ความรู้ 5. สรุปและ ประเมินค่าการหาคำตอบและ 6. นำเสนอและประเมินผลงาน



ภาพที่ 3: รูปแบบกิจกรรมการพัฒนาทักษะการออกแบบขั้นตอนวิธีทางคอมพิวเตอร์ โดยใช้เทคโนโลยีเว็บ 2.0 ด้วยเทคนิคพี่เลี้ยง

3) ระยะที่ 3 ติดตามผล เป็นกระบวนการในติดตามผลการคิดอย่างเป็นระบบของผู้เรียนจากการใช้รูปแบบการพัฒนาทักษะ โดยจะดำเนินการติดตามจากผู้เรียน และอาจารย์ผู้สอน

ตารางที่ 4: ความคิดเห็นของผู้เชี่ยวชาญที่มีต่อรูปแบบกิจกรรม

รายการ	ระดับความคิดเห็น		
	\bar{X}	S.D.	ความหมาย
ความเหมาะสมขององค์ประกอบของรูปแบบกิจกรรม	4.57	0.51	มากที่สุด
ระยะที่ 1 เตรียมความพร้อม	4.68	0.46	มากที่สุด
ระยะที่ 2 พัฒนาทักษะ	4.48	0.52	มาก
ระยะที่ 3 ติดตามผล	4.56	0.57	มากที่สุด

4.2.3 ผลการศึกษาความคิดเห็นของผู้เชี่ยวชาญที่มีต่อรูปแบบกิจกรรมการพัฒนาทักษะการออกแบบขั้นตอนวิธีทางคอมพิวเตอร์ โดยใช้เทคโนโลยีเว็บ 2.0 ด้วยเทคนิคพี่เลี้ยง ดังแสดงในตารางที่ 4

จากตารางที่ 4 พบว่า ความคิดเห็นของผู้เชี่ยวชาญที่มีต่อความเหมาะสมขององค์ประกอบของรูปแบบกิจกรรมอยู่ในระดับมากที่สุด ($\bar{X}=4.57$, S.D.=0.51) และเมื่อพิจารณาเป็นระยะ พบว่าระยะที่ 1 และระยะที่ 2 มีความเหมาะสมในระดับมากที่สุด คือระยะการเตรียมความพร้อม ($\bar{X}=4.68$, S.D.=0.46) ระยะการติดตามผล ($\bar{X}=4.56$, S.D.=0.57) ตามลำดับ และระยะการพัฒนาทักษะมีความเหมาะสมในระดับมาก ($\bar{X}=4.48$, S.D.=0.52)

5. สรุปผลการวิจัย

การวิจัยในครั้งนี้ พบว่า รูปแบบการพัฒนาทักษะการออกแบบขั้นตอนวิธีทางคอมพิวเตอร์ มีองค์ประกอบ 4 ส่วน คือ 1) หลักการที่เกี่ยวข้อง 2) แนวคิด ทฤษฎี 3) การพัฒนาทักษะ และ 4) เทคโนโลยีสารสนเทศและการสื่อสาร โดยความคิดเห็นของผู้เชี่ยวชาญที่มีต่อองค์ประกอบของรูปแบบโดยรวมอยู่ในระดับมากที่สุด ($\bar{X}=4.73$, S.D.=0.45) และรูปแบบกิจกรรมการพัฒนาทักษะการออกแบบขั้นตอนวิธีทางคอมพิวเตอร์ โดยใช้เทคโนโลยีเว็บ 2.0 ด้วยเทคนิคพี่เลี้ยง มีกิจกรรมการดำเนินงาน 3 ระยะ คือ ระยะที่ 1 การเตรียมความพร้อม ระยะที่ 2 การพัฒนาทักษะ และ ระยะที่ 3 การติดตามผล โดยความคิดเห็นของผู้เชี่ยวชาญที่มีต่อกิจกรรมโดยรวมอยู่ในระดับมากที่สุด ($\bar{X}=4.57$, S.D.=0.51)

เอกสารอ้างอิง

- [1] วิทยา อธิราษฎร์, การออกแบบขั้นตอนวิธีทางคอมพิวเตอร์, มหาสารคาม : คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏมหาสารคาม, 2548.
- [2] สายชล จินใจ, การพัฒนารูปแบบการเรียนการสอนแบบผสมผสานรายวิชาการเรียนโปรแกรมภาษาคอมพิวเตอร์ 1 สาขาคอมพิวเตอร์ธุรกิจ, ปรัชญาดุษฎีบัณฑิตสาขาคอมพิวเตอร์ศึกษา บัณฑิตวิทยาลัยสถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2550.
- [3] สมชาย พัฒนาชวนชม, การเขียนโปรแกรมสำหรับเด็กเวอร์ชัน 1.0, ศรีเอทีฟคอมมอนส์ 3.0, 2555.
- [4] ทิศนา แจมมณี, รูปแบบการเรียนการสอนของการเรียนรู้แบบร่วมมือ (Instructional Models of cooperative Learning), กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2548.

- [5] พิสุทธา อารีราษฎร์, *โครงสร้างข้อมูล*, มหาสารคาม : คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏมหาสารคาม, 2542.
- [6] เกษม ไคร้มา, “ระบบพี่เลี้ยง (Mentoring) ในการจัดการความรู้” *บทความออนไลน์* [ออนไลน์]. เข้าถึงได้จาก <http://www.e-manage.mju.ac.th/articleDetail.aspx?qid=136&> (เมื่อ 1 ธันวาคม 2554)
- [7] มัณฑรา ธรรมบุศย์, “การพัฒนาคุณภาพการเรียนรู้โดยใช้ PBL” *บทความออนไลน์* [ออนไลน์]. เข้าถึงได้จาก <http://www.edu.chandra.ac.th/teacherAll/mdra/dara/PBL.doc> (เมื่อ 29 มิถุนายน 2553)
- [8] จรีพร วีระพันธุ์, *การออกแบบและวิเคราะห์ขั้นตอนวิธีทางคอมพิวเตอร์*, กรุงเทพฯ: จุฬาลงกรณ์มหาวิทยาลัย, 2552.

กลไกค้นหาเกตเวย์ที่มีโอเวอร์เฮดจากการแลกเปลี่ยนข้อความต่ำ สำหรับเครือข่ายเคลื่อนที่เฉพาะกิจ

Gateway Discovery Mechanism with Low Message Exchange Overhead for Mobile Ad Hoc Networks

สุเมธ ประภาวัต(Sumet Prabhavat)¹ สารัช ศรีสวัสดิ์พงษ์ (Sarach Sriswadpong)²

ทะนง ชูวนิชชานนท์ (Tanong Chuvanitchanon)³ วรวัชร ฌรณกะชวณะ(Worrawat Narongkhachavana)⁴

คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

¹sumet@it.kmitl.ac.th, ²acecharmil@gmail.com, ³tanong3200@hotmail.com, ⁴worravat@gmail.com

บทคัดย่อ

ในการค้นหาเกตเวย์เพื่อให้เครือข่ายเคลื่อนที่เฉพาะกิจติดต่อกับเครือข่ายภายนอกหรืออินเทอร์เน็ตได้ โอเวอร์เฮดที่เกิดจากการแลกเปลี่ยนข้อมูลระหว่างโหนดเพื่อการค้นหาเกตเวย์ และเวลาที่ต้องใช้ในการค้นหาเป็นปัจจัยที่ลดทอนประสิทธิภาพเครือข่ายอย่างหลีกเลี่ยงไม่ได้ ในงานวิจัยก่อนหน้านี้พยายามแก้ไขปัญหาดังกล่าว เช่น กลไกในโพรโทคอล AODV ที่ใช้เวลาในการค้นหาเกตเวย์น้อย แต่ก็ทำให้เกิดโอเวอร์เฮดมาก เมื่อเทียบกับกลไกในโพรโทคอล OLSR ที่มีโอเวอร์เฮดต่ำ แต่ก็ใช้เวลาในการค้นหานาน บทความนี้จึงเสนอกลไกที่สามารถค้นหาเกตเวย์ได้โดยมีโอเวอร์เฮดต่ำ และใช้เวลาในการค้นหาน้อย ซึ่งพิสูจน์ได้โดยการจำลองการทำงานและเปรียบเทียบกับกลไกค้นหาเกตเวย์จากงานวิจัยก่อนหน้าในสถานะที่โหนดเคลื่อนที่แบบสุ่ม และมีพารามิเตอร์การจำลองที่แตกต่างกันออกไป

คำสำคัญ: เครือข่ายเคลื่อนที่เฉพาะกิจ การค้นหาเกตเวย์ โอเวอร์เฮด การแลกเปลี่ยนข้อมูลเพื่อหาเกตเวย์

Abstract

A gateway discovery mechanism on mobile ad hoc network (MANET) is applicable for connecting MANET to other networks or the Internet. However, it causes message exchange overhead and gateway discovery time (leading to network delay), which degrade network performance significantly. These issues have been addressed by previous works: A gateway discovery mechanism in a modified AODV takes short time to find a gateway but causes high overhead whereas that in a modified version of OLSR needs only a few number of

message exchanges but requires long discovery time. This paper proposes an improved gateway discovery mechanism that can reduce message exchange overhead while maintains low gateway discovery time. Performance evaluations are conducted using simulations with different parameters and random waypoint mobility model.

Keywords: Mobile Ad Hoc Network Gateway Discovery Message Exchange Overhead

1. บทนำ

จากเทคโนโลยีด้านการสื่อสารที่มีการพัฒนาไปอย่างรวดเร็วอันส่งผลให้เกิดนวัตกรรมใหม่ๆ เกิดขึ้น เช่น การสื่อสารแบบไร้สาย และมีการใช้งานกันอย่างแพร่หลายซึ่งการสื่อสารกลายเป็นสิ่งสำคัญในการดำเนินชีวิตของมนุษย์ไปแล้ว ทั้งนี้หากเกิดสถานการณ์ที่เกิดภัยพิบัติหรือเหตุการณ์ไม่คาดฝันซึ่งทำให้การติดต่อสื่อสารระหว่างโหนดกับโครงสร้างพื้นฐานถูกตัดขาด ส่งผลให้ไม่สามารถติดต่อสื่อสารและขอความช่วยเหลือจากบุคคลอื่น ๆ ได้ จากปัญหาดังกล่าวจึงมีนักวิจัยหลายกลุ่มพยายามพัฒนาเครือข่ายสื่อสารเฉพาะกิจของโหนดเคลื่อนที่ (MANET : Mobile Ad-hoc Network) [1] ซึ่งทำให้โหนดเคลื่อนที่สามารถติดต่อสื่อสารกันได้โดยไม่ต้องพึ่งพาโครงสร้างพื้นฐาน

แต่การเชื่อมต่อกันเพียงในเครือข่ายเฉพาะกิจดังกล่าวทำให้เกิดปัญหาการจำกัดทางด้านขอบเขตของเครือข่าย และจากที่กล่าวมาการที่โหนดเคลื่อนที่ทำหน้าที่เป็นผู้ส่งต่อข้อมูลให้โหนดอื่น ๆ นั้นทำให้มีการใช้พลังงานสำหรับทำหน้าที่ดังกล่าวออกเหนือจากเพื่อการส่งข้อมูลของตนเอง อันส่งผลให้

เกิดปัญหาในการใช้พลังงานที่มีอยู่อย่างจำกัด (จากแบตเตอรี่) ของโมบายล์โหนด

จากปัญหาที่กล่าวมาผู้จัดทำจึงศึกษารูปแบบการหาเส้นทางให้สามารถหาเส้นทางเชื่อมต่อที่ผ่านเกตเวย์ออกไปยังโครงสร้างพื้นฐานได้บนเครือข่ายเฉพาะกิจของโมบายล์โหนด (อย่างเช่น เครือข่ายอินเทอร์เน็ต) เพื่อแก้ไขปัญหาด้านการจำกัดขอบเขตเครือข่าย ซึ่งทำให้ติดต่อสื่อสารออกไปนอกพื้นที่ได้ และลดการใช้พลังงานที่มีอยู่อย่างจำกัดของโมบายล์โหนดให้ทำงานได้นานขึ้น ไปจนกระทั่งลดโอกาสเกิดความไม่ต่อเนื่องในเครือข่าย

2. ทฤษฎีและงานที่เกี่ยวข้อง

โพรโทคอลค้นหาเส้นทาง (Routing Protocols) ในเครือข่ายสื่อสารเฉพาะกิจถูกแบ่งเป็น 3 กลุ่ม [2] คือ 1) โพรโทคอลค้นหาเส้นทางแบบรีแอกทีฟ (Reactive Routing Protocols) ซึ่งค้นหาเส้นทางเมื่อต้องการส่งข้อมูล เช่น โพรโทคอลเอโอดีวี (AODV : Ad hoc On-Demand Distance Vector) [4] และ โพรโทคอลดีเอสอาร์ (DSR : Dynamic Sort Routing) [3] 2) โพรโทคอลค้นหาเส้นทางแบบโพรแอกทีฟ (Proactive Routing Protocols) ทำการค้นหาเส้นทางไปยังทุกปลายทางที่เป็นไปได้ก่อนที่จะมีการส่งข้อมูล เช่น โพรโทคอลโอแอลเอสอาร์ (OLSR : Optimized Link State Routing) [5] และสุดท้าย 3) โพรโทคอลค้นหาเส้นทางแบบไฮบริดจ์ (Hybrid Routing Protocols) ซึ่งมีการทำงานผสมผสานระหว่างโพรโทคอลค้นหาเส้นทางทั้ง 2 แบบก่อนหน้า

โดยมีงานวิจัยที่ได้ศึกษาการเชื่อมต่อระหว่างเครือข่ายเฉพาะกิจของโมบายล์โหนดกับเครือข่ายอื่น ๆ เช่น เครือข่ายโทรศัพท์ (Cellular Network) [10] เครือข่ายอินเทอร์เน็ต [6],[7],[8] โดยใช้การค้นหาโหนดเกตเวย์ในการเชื่อมต่อระหว่างเครือข่ายสื่อสารเฉพาะกิจและเครือข่ายอินเทอร์เน็ตหรือเครือข่ายอื่น ๆ ซึ่งเกตเวย์มีหน้าที่เป็นตัวกลางในการเชื่อมต่อ

จากการศึกษางานวิจัยที่เกี่ยวข้องกับการค้นหาเกตเวย์ มีการนำโพรโทคอลค้นหาเส้นทาง AODV [6],[8] มาปรับแก้รูปแบบและการส่งของข้อความ (Message) จากที่กำหนดใน RFC [4] โดยแก้ไข HELLO messages ที่ใช้เพื่อบอกถึงการมีตัวตนของตนเองกับโหนดรอบข้าง โดยการ broadcast ซึ่งถูก

ส่งก็ต่อเมื่อโหนดนั้น ๆ ยังไม่มีการส่งคำร้องขอเส้นทาง (Route Request) การตอบกลับเส้นทาง (Route Reply) โดยในการปรับแก้รูปแบบข้อความนั้นได้ทำการเพิ่มแฟล็ก (Flag) พิเศษ ที่เรียกว่าแฟล็ก I (I-flag) เพื่อระบุว่าเป็น HELLO message ที่ส่งมาจากเกตเวย์ และบันทึกแอดเดรสของเกตเวย์ (Gateway Address) ใน HELLO message ด้วยเพื่อระบุว่าคุณสร้างมาจากเกตเวย์โหนดใด ขั้นตอนในการส่ง HELLO_I นั้น จะเริ่มส่งจากโหนดเกตเวย์เพื่อบอกถึงการมีอยู่ของโหนดเกตเวย์ให้ทุกโหนดที่อยู่ภายในช่วงคลื่นสัญญาณของเกตเวย์รับรู้ผ่านการรับ HELLO_I (จากการ broadcast HELLO_I ของเกตเวย์) จากนั้นโหนดที่ได้รับ HELLO_I จะทำการการบันทึกเส้นทางที่สามารถไปยังเกตเวย์ และ broadcast HELLO_I (Rebroadcast) HELLO_I ซ้ำออกไป แต่ถ้าโหนดใดไม่ได้รับ HELLO_I ในช่วงระยะเวลาของการส่ง HELLO messages (HELLO interval) โหนดนั้นก็ทำการส่ง HELLO messages แบบปกติที่ไม่มีการปรับแก้แทน ซึ่งขอเรียกการค้นหาเส้นทางไปยังเกตเวย์นี้ว่า อัลกอริทึม HELLO และภายในงานวิจัยดังกล่าวนักวิจัยได้ทำการเปรียบเทียบอัลกอริทึม HELLO กับอัลกอริทึมค้นหาเกตเวย์แบบโพรแอกทีฟ และแบบรีแอกทีฟ โดยผลสรุปคือ อัลกอริทึม HELLO มีเวลาที่ใช้ในการค้นหาเกตเวย์ที่น้อยกว่าเมื่อเปรียบเทียบกับอัลกอริทึมทั้งสอง

โดยสำหรับโพรโทคอลค้นหาเส้นทางแบบโพรแอกทีฟ [8] เมื่อโหนดได้รับข้อมูลเกี่ยวกับเกตเวย์จะตัดสินใจว่าจะรับหรือทิ้งข้อมูลนั้น โดยจะรับข้อมูลนั้นก็ต่อเมื่อโหนดไม่รู้เส้นทางไปยังเกตเวย์หรือข้อมูลนั้นมีเส้นทางที่สั้นกว่าหรือข้อมูลนั้นมาจากแหล่งกำเนิดเดียวกันแต่มีความใหม่กว่า (มี sequence number ที่มากกว่า) และถ้าหากโหนดเลือกรับข้อมูลเส้นทางก็จะหยุดการร้องขอเส้นทาง จากนั้นจึงทำการตั้งค่าหรือปรับปรุงเส้นทางไปยังเกตเวย์ตามข้อมูลที่ได้รับมา และส่งต่อไปยังโหนดอื่น ๆ

3. แนวคิดและการดำเนินงาน

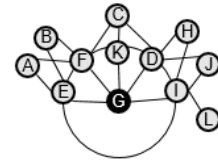
จากงานวิจัยในส่วนที่ 2 เห็นได้ว่าอัลกอริทึม HELLO มีเวลาในการทำให้แต่ละโหนดรับรู้เส้นทางเชื่อมต่อไปยังเกตเวย์น้อยที่สุดเมื่อเทียบกับอัลกอริทึมอื่นๆ [8] แต่ยังมีข้อเสียเรื่องโอเวอร์เฮดเพราะทุกโหนดต้องทำการ broadcast ข้อมูล

```

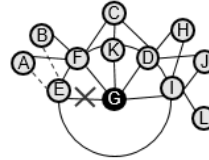
1:   for x in V
2:     assume N = Nx - Nx'
3:     for i in N
4:       for j in N - i
5:         if Nj ⊆ Ni
6:           N = N - j
7:         end
8:       end
9:     end
10:    for i in N
11:      for j in N - i
12:        Ni,j = Ni,j ∪ (Ni ∩ Nj)
13:      end
14:    end
15:    for i in N
16:      if Ni ⊆ Ni,j
17:        N = N - i
18:        Ni,j = Ni,j - Ni
19:      end
20:    end
21:    V = V - (Nx - N)
22:  end
    
```

ภาพที่ 1: แสดงโค้ดเทียมของอัลกอริทึมในการหาโหนด MPR

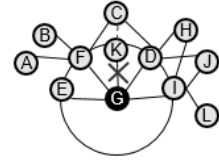
ในงานวิจัยนี้จึงนำอัลกอริทึม HELLO มาเพิ่มกลไกเพื่อลดจำนวนการบรอดแคสต์ข้อมูลเส้นทางที่เกี่ยวข้องกับเกตเวย์ ด้วยวิธีการลดจำนวนโหนดที่ทำการบรอดแคสต์ ซึ่งกลไกนี้ถูกใช้ในโพรโทคอลค้นหาเส้นทาง OLSR เรียกว่าการเลือกโหนดเอ็มพัวร์ (MPR: Multipoint Relays Selection) โดยมีหน้าที่ในการส่งต่อข้อมูล (Retransmission) โดยโหนดที่เริ่มต้นบรอดแคสต์นั้นจะมีทำหน้าที่ในการคัดเลือกโหนด MPR หรือเป็น MPR Selector ซึ่งโหนด MPR จะอยู่ในกลุ่มของโหนดข้างเคียง (Neighbor nodes) แต่ทุกโหนดที่เป็นโหนดข้างเคียงไม่จำเป็นต้องเป็นโหนด MPR ซึ่งสามารถเปลี่ยนแปลงได้ตามความเหมาะสมในการคัดเลือกโหนด MPR นั้นโหนด MPR Selector จะเลือกจากกลุ่มโหนดข้างเคียงในระยะ 1 hop (One Hop Neighbors) ของตนที่สามารถเชื่อมต่อกับโหนดข้างเคียงในระยะ 2 hop (Two Hop Neighbors) ของตนที่มากที่สุด ซึ่งโหนดข้างเคียงในระยะ 1 hop คือ โหนดที่ตนเองสามารถเชื่อมต่อได้ (โหนดที่อยู่ภายในช่วงคลื่นสัญญาณของตนเอง) แต่โหนดข้างเคียงในระยะ 2 hop คือ โหนดที่ตนเองไม่สามารถเชื่อมต่อได้แต่สามารถเชื่อมต่อผ่านโหนดข้างเคียงในระยะ 1 hop ได้ โดยมีงานวิจัยบางส่วนได้นำเอากลไกการเลือกโหนด MPR มาประยุกต์เข้ากับกลไกการแลกเปลี่ยน HELLO messages บนโพรโทคอล



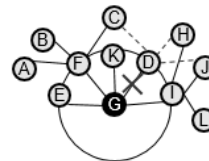
ก. การกระจาย HELLO_I message ของอัลกอริทึม HELLO



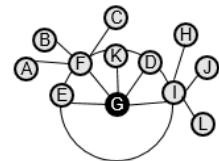
ข. ลบ E ออกจากกลุ่ม MPR



ค. ลบ K ออกจากกลุ่ม MPR



ง. ลบ D ออกจากกลุ่ม MPR



จ. กลุ่ม MPR คือ F และ I

ภาพที่ 2: แสดงตัวอย่างการหาโหนด MPR

AODV [11] ซึ่งมีความแตกต่างกับงานวิจัยของเราเนื่องจากผู้จัดทำได้ทำการออกแบบกลไกการคัดเลือกโหนด MPR และกลไกการทำงานของโหนด MPR ขึ้นเอง ซึ่งแต่ละกลไกมีการทำงานดังนี้ กลไกการคัดเลือกโหนด MPR จะทำโดยการแลกเปลี่ยน HELLO message เพื่อเรียนรู้โหนดข้างเคียงในระยะ 1 hop และโหนดข้างเคียงในระยะ 2 hop ของทุกโหนดข้างเคียงในระยะ 1 hop จากนั้นจะนำข้อมูลดังกล่าวที่เรียนรู้มาทำการคัดเลือกโหนด MPR ซึ่งมีการทำงานตามโค้ดเทียมในภาพที่ 1

- โดย $G =$ กราฟเชื่อมโยง โดยที่ $G = (V, E)$
- $V =$ เซ็ตของโหนด
- $E =$ เส้นทางเชื่อมโยงของแต่ละโหนด
- $x' =$ โหนด MPR ก่อนหน้า (Predecessor node)
- $N_x \in V =$ เซ็ตของโหนดข้างเคียงในระยะ 1 hop (1-hop neighbors) ของ x ใดๆใน V
- $N'_x \in V =$ เซ็ตของโหนดข้างเคียงในระยะ 1 hop ของ x ใดๆที่ไม่มีโหนดที่ก่อให้เกิดเส้นทางที่ซ้ำซ้อน (Redundancy Paths) โดยที่ $N'_x \in V$ เป็นเซตย่อยของ $N_x \in V$ และทำให้ $G' = (V', E')$ โดยที่ $V' = U(x \in V) N'_x$ หรือกล่าวได้ว่า $N'_x \in V$ คือกลุ่มของโหนด MPR

$N_{i,j}$ = เซ็ตของโหนดข้างเคียงในระยะ 1 hop (1-hop neighbors) ร่วมของโหนด i และ j โดย $i, j \in N$ และ $i \neq j$ จากโค้ดเทียมในภาพที่ 1 แสดงถึงอัลกอริทึมที่ใช้ในการคำนวณหาโหนด MPR หรือ $N^*x \in V$ และภาพที่ 2 แสดงตัวอย่างการหาโหนด MPR โดยการทำงานของอัลกอริทึมแบ่งออกได้เป็น 2 ขั้นตอน 1) หากความเป็นเซตย่อย (Subset) คือ หาเซตของโหนดข้างเคียงในระยะ 1 hop ของโหนดข้างเคียงที่เหลือที่ครอบคลุมโหนดข้างเคียงในระดับ 2 hop ที่เหมือนกันหรือน้อยกว่า แสดงอยู่ในภาพที่ 1 บรรทัดที่ 3 ถึง 9 และภาพที่ 2(ข และ ค) และ 2) หาเซตของโหนดข้างเคียงในระยะ 1 hop ของโหนดข้างเคียงที่เหลือที่ครอบคลุมโหนดข้างเคียงในระดับ 2 hop ของตนมากที่สุด แสดงอยู่ในภาพที่ 1 บรรทัดที่ 10 ถึง 20 และภาพที่ 2(ง) โดยทำขั้นตอนที่ 2) ซ้ำจนกว่าจะไม่สามารถตัดเซตใดออกจากการพิจารณาได้อีกซึ่งจะผลเป็นเซตของโหนด MPR ซึ่งแสดงให้ภาพที่ 2(จ)

และสำหรับกลไกการทำงานของโหนด MPR จะระบุให้โหนดที่กำลังส่งข้อมูลเกตเวย์ขณะนั้นจะเป็นผู้เลือกโหนด MPR โดยใช้กลไกการเลือกโหนด MPR ที่ได้กล่าวไว้ข้างต้น ดังนั้นคำนวณหาโหนด MPR นั้นจึงเริ่มต้นคำนวณจากโหนดเกตเวย์เมื่อคำนวณเสร็จเกตเวย์จะทำการส่ง HELLO_I message ที่บรรจุข้อมูลเพื่อระบุว่าโหนดใดเป็นโหนด MPR ไปยังโหนดข้างเคียงในระยะ 1 hop จากนั้นเมื่อโหนดข้างเคียงที่เป็น MPR ได้รับ HELLO_I message ดังกล่าวก็จะทำการคำนวณหาโหนด MPR ของตนต่อไป และส่ง HELLO_I message ไปยังโหนด MPR ที่คำนวณได้ต่อไปเรื่อยๆ พร้อมทั้งบันทึกเส้นทางไปยังเกตเวย์ตามอัลกอริทึม HELLO ที่ได้กล่าวไว้ในส่วนที่ 2 แต่สำหรับโหนดข้างเคียงที่ได้รับ HELLO_I message ไม่ใช่โหนด MPR ก็จะไม่ส่งต่อ HELLO_I message แต่บันทึกเส้นทางไปยังเกตเวย์เท่านั้น

ดังนั้นอัลกอริทึมที่ได้พัฒนาได้มีการลดโอเวอร์เฮดจากการแลกเปลี่ยน HELLO_I message ของอัลกอริทึม HELLO คือทุกโหนดที่ทำหน้าที่บรอดแคสต์ HELLO_I message ซึ่งอาจทำให้เกิดความซ้ำซ้อน ดังภาพที่ 2(ก) ที่ทุกโหนดต้องทำหน้าที่บรอดแคสต์ข้อความ แต่สำหรับอัลกอริทึมที่ได้พัฒนาโดยขอเรียกว่า HELLOMPR จะมีเพียงบางโหนดที่ทำหน้าที่ในการบรอดแคสต์ ตัวอย่างเช่น ในภาพที่ 2(ก) ที่ใช้อัลกอริทึม

HELLO ที่โหนด A และ B ได้รับ HELLO_I message จากทั้งโหนด E และ F เมื่อเปรียบเทียบกับภาพที่ 2(จ) ที่ใช้อัลกอริทึม HELLOMPR ที่โหนด A และ B จะได้รับ HELLO_I message จากโหนด F เท่านั้น

4. การทดลองและการประเมินผล

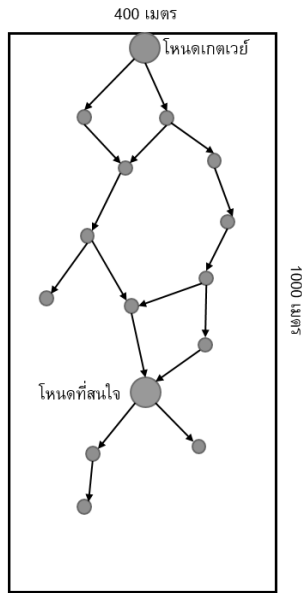
ในการทดลองได้ทำการจำลองการทำงานของอัลกอริทึมผ่านโปรแกรมจำลอง OMNET++ ซึ่งเป็นการเปรียบเทียบระหว่างอัลกอริทึมในการค้นหาเกตเวย์ทั้ง 3 ได้แก่ 1) อัลกอริทึมโพรแอกทีฟ 2) อัลกอริทึม HELLO และ 3) อัลกอริทึม HELLOMPR ซึ่งในอัลกอริทึมที่ 1) และ 2) อ้างอิงการทำงานตามงานวิจัยที่ [6] โดยมีข้อกำหนดเกี่ยวกับเหตุการณ์จำลอง และผลการทดลองเปรียบเทียบดังนี้

4.1 เหตุการณ์จำลองและตัวแปรต่าง ๆ ที่ใช้ในการจำลอง

กำหนดให้การกระจายตัวของโหนดจะกระจายแบบสุ่มอยู่ในเครือข่ายพื้นที่ขนาดกว้าง 400 เมตร และยาว 1,000 เมตร ดังแสดงในภาพที่ 3 ซึ่งแต่ละโหนดจะเคลื่อนที่ภายใต้ความเร็วสูงสุด 10 เมตร / วินาที โดยในการทดลองแต่ละครั้งใช้การกระจายตัวของโหนดแบบสม่ำเสมอ (Uniform Distribution) เพื่อให้แต่ละการทดลองโหนดมีการกระจายตัวที่แตกต่างกัน โดยจากภาพที่ 3 กำหนดให้โหนดที่เป็นเกตเวย์ (Gateway node) อยู่บนสุดของเครือข่าย และกำหนดให้โหนดที่สนใจซึ่งคือโหนดที่จะต้องได้รับข้อมูลเส้นทางไปยังเกตเวย์อยู่ห่างจากเกตเวย์ระยะ 600 เมตร โดยที่ทุกโหนดมีระยะในการส่งข้อมูล 250 เมตรรอบตัวโหนด และมีตัวแปรต่าง ๆ ที่ได้กำหนดและใช้ในการจำลองเพิ่มเติมดังแสดงในตารางที่ 1

ตารางที่ 1: ตัวแปรต่างๆที่กำหนดในการทดลอง

จำนวนโหนดในเครือข่ายสื่อสารเฉพาะกิจ	15, 60
พื้นที่ของเครือข่าย	400 เมตร × 1,000 เมตร
ระยะส่งข้อมูลของโหนด	250 เมตร
เวลาทั้งหมดในการทดลอง	300 วินาที
ช่วงเวลาในการส่ง HELLO message	1, 5, 10, 15, 20, 25, 30
จำนวนรอบที่ใช้ในการทดลอง	100 รอบ (ทั้ง 15,60 โหนด)
รูปแบบการเคลื่อนที่ของโหนด	Random Waypoint
ความเร็วสูงสุดในการเคลื่อนที่	10 เมตร/วินาที



ภาพที่ 3: แสดงภาพจำลองทอพอโลยีของเครือข่าย

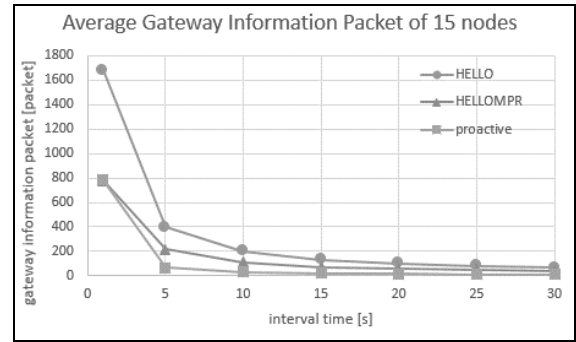
4.2 ผลการทดลองและตัวชี้วัด

สำหรับการทดลองเพื่อเปรียบเทียบระหว่างอัลกอริทึมค้นหาเกตเวย์เราได้เลือกตัวชี้วัดในการทดลองคือ 1) จำนวนของข้อความที่บรรจุข้อมูลเส้นทางไปยังเกตเวย์ที่ถูกส่งภายในระบบ 2) เวลาที่ใช้ในการค้นหาเกตเวย์

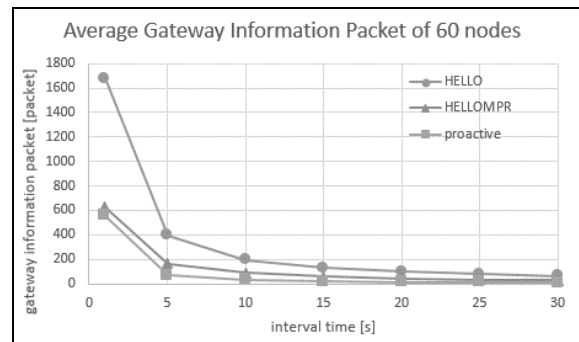
ในภาพที่ 4 ภาพที่ 5 และตารางที่ 2 แสดงให้เห็นถึงผลการทดลองในตัวชี้วัดด้านจำนวนข้อความที่บรรจุข้อมูลเส้นทางไปยังเกตเวย์ โดยอัลกอริทึม HELLO มีจำนวนข้อความที่ส่งในเครือข่ายมากที่สุด เนื่องจากทุกโหนดต้องทำการบรอดแคสต์ข้อมูลของเกตเวย์ แต่อัลกอริทึม HELLOMPR มีจำนวนน้อยกว่าในอัลกอริทึม HELLO มากเนื่องจากมีกลไกในการลดจำนวนโหนดที่ทำการบรอดแคสต์ แต่อัลกอริทึมแบบโพรแอกทีฟก็ยังมีจำนวนข้อความน้อยที่สุดเนื่องมาจากทุกโหนดจะทราบ topology ของทั้งระบบทำให้สามารถคำนวณหาโหนด MPR ได้อย่างเต็มประสิทธิภาพแต่สำหรับอัลกอริทึมที่ได้พัฒนานั้นทราบเพียงโหนดข้างเคียงในระยะ 1 และ 2 hop

ตารางที่ 2: ผลลัพธ์จำนวนข้อความที่ส่งเพื่อการค้นหาเกตเวย์

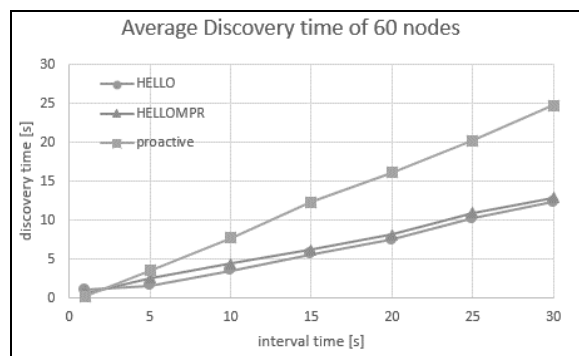
	Interval [s]		
	1	10	30
15 nodes			
proactive	784.443 pkt	31.402 pkt	10 pkt
HELLO	1683.25 pkt	200.39 pkt	66.124 pkt
HELLOMPR	778.248 pkt	110.69 pkt	37.009 pkt
60 nodes			
proactive	558.348 pkt	31.384 pkt	10 pkt
HELLO	1680.24 pkt	192.42 pkt	65.234 pkt
HELLOMPR	628.518 pkt	87.446 pkt	27.771 pkt



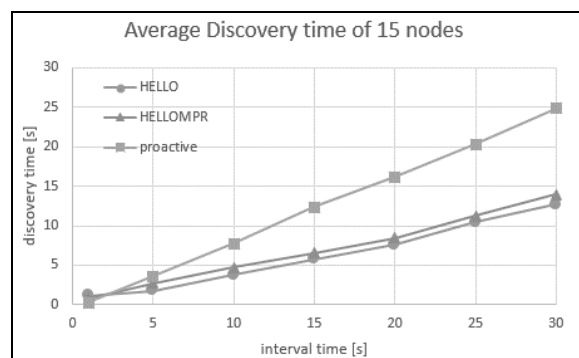
ภาพที่ 4: กราฟจำนวนข้อความที่ใช้เพื่อค้นหาเกตเวย์สำหรับ 15 โหนด



ภาพที่ 5: กราฟจำนวนข้อความที่ใช้เพื่อค้นหาเกตเวย์สำหรับ 60 โหนด



ภาพที่ 6: กราฟเวลาที่ใช้ในการค้นหาเกตเวย์สำหรับจำนวน 15 โหนด



ภาพที่ 7: กราฟเวลาที่ใช้ในการค้นหาเกตเวย์สำหรับจำนวน 60 โหนด

ตารางที่ 3: ผลลัพธ์เวลาเฉลี่ยที่ใช้ในการค้นหาเกตเวย์

15 nodes	Interval [s]		
	1	10	30
proactive	0.231s	7.741s	24.83s
HELLO	1.148s	3.757s	12.69s
HELLOMPR	0.717s	4.677s	13.91s
60 nodes			
proactive	0.221s	7.727s	24.74s
HELLO	1.015s	3.532s	12.34s
HELLOMPR	0.621s	4.391s	12.89s

ในภาพที่ 6 ภาพที่ 7 และตารางที่ 3 แสดงผลลัพธ์ในตัวชี้วัดด้านเวลาที่ใช้ในการค้นหาเกตเวย์ (Gateway Discovery Time) โดยจะเริ่มนับตั้งแต่เกตเวย์ประกาศข้อมูลเส้นทางจนกระทั่งโหนดที่สนใจรับรู้เส้นทางไปยังเกตเวย์ โดยอัลกอริทึมแบบโพรแอกทีฟใช้เวลาในการค้นหาเกตเวย์สูง แต่สำหรับอัลกอริทึม HELLO และอัลกอริทึม HELLOMPR ก็มีเวลาที่ใช้ในการค้นหาเกตเวย์เพียงเล็กน้อย แต่อัลกอริทึม HELLO จะใช้เวลาในน้อยที่สุดเนื่องจากการเพิ่มกลไกคัดเลือกโหนด MPR ทำให้โหนดที่สนใจอาจได้รับ HELLO_I message ซ้ำกว่าในกรณีที่โหนดที่สนใจนั้นอยู่ใกล้กับโหนดไม่ใช่ MPR มากกว่าโหนดที่ทำหน้าที่เป็น MPR

5. สรุปผล

ในงานวิจัยนี้ได้พัฒนาอัลกอริทึมในการค้นหาเกตเวย์สำหรับเครือข่ายสื่อสารเฉพาะกิจโดยมีพื้นฐานมาจากโพรโทคอล AODV และได้นำหลักการในการคัดเลือกโหนด MPR จากโพรโทคอล OLSR โดยใช้ทฤษฎีเซตมาช่วยในการคำนวณเพื่อลดจำนวนข้อมูลที่ส่งต่อภายในเครือข่าย

จากผลการทดลองสรุปได้ว่าเครือข่ายที่ใช้อัลกอริทึม HELLOMPR มีการส่งข้อมูลเกี่ยวกับเกตเวย์น้อยกว่าอัลกอริทึม HELLO จนเข้าใกล้อัลกอริทึมแบบโพรแอกทีฟ ซึ่งเป็นผลจากการลดจำนวนโหนดที่ต้อง broadcast ข้อมูลที่ใช้ในการค้นหาเกตเวย์ ซึ่งส่งผลให้โหนดในเครือข่ายมีพลังงานทำให้โหนดอยู่ในระบบได้นานขึ้น และประสิทธิภาพด้านเวลาที่ใช้ในการทำให้โหนดที่สนใจรู้จักเส้นทางไปยังเกตเวย์ไม่ได้ลดลง ซึ่งในสถานการณ์ภัยพิบัตินั้นมีความจำกัดด้านแหล่งพลังงานของโบายล์โหนด (จากแบตเตอรี่) และความเร่งด่วนในการสื่อสารนั้นก็มีความสำคัญ ดังนั้นอัลกอริทึมที่มีความเหมาะสมคืออัลกอริทึม HELLOMPR เพราะมีการส่งข้อความเป็นจำนวนน้อย และมีเวลาในการค้นหาเกตเวย์ต่ำ

เอกสารอ้างอิง

- [1] T. S. Rappaport. "Wireless Communications Principles & Practice." Upper Saddle River, New Jersey: Prentice-Hall, Inc.1996
- [2] T. Lin. "Mobile Ad-hoc Network Routing Protocols: Methodologies and Applications." Dissertation submitted to the Faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the Ph.D. in Computer Engineering. 2004
- [3] D. B. Johnson, D. A. Maltz, Y. C. Hu, and J. G. Jetcheva, "The Dynamic Source Routing Protocol for Mobile Ad Hoc Networks (DSR)." Internet Engineering Task Force (IETF) draft, February 2007.
- [4] C. E. Perkins, E. M. Belding-Royer, and S. R. Das, "Ad hoc On-Demand Distance Vector(AODV) Routing." Internet Engineering Task Force (IETF) draft, July 2003.
- [5] T. Clausen, P. Jacquet, A. Laouiti, P. Minet, P. Muhlethaler, A. Qayyum, and L. Viennot, "Optimized Link State Routing Protocol (OLSR)." Internet Engineering Task Force(IETF) draft, October 2003.
- [6] M. Rosenschon. "Internet Gateway Discovery for Mobile Ad Hoc Networks" Submitted for the Degree of Doctor of Philosophy School of Engineering and Mathematical Sciences CITY University London. 2 October 2007
- [7] J. Broch, et al. "Supporting Hierarchy and Heterogeneous Interfaces in Multi-Hop Wireless Ad Hoc Networks" International Symposium on Parallel Architectures, Algorithms and Networks (ISPA99), Workshop on Mobile Computing, Perth, Western Australia. June 1999
- [8] M. Rosenschon, T. Manz, J. Habermann, V. Rakocevic. "Gateway Discovery Algorithm for Ad-Hoc networks using HELLO Messages" International Workshop on Wireless Ad Hoc Networks IWWAN 2005, London. May 2005.
- [9] M. Rosenschon, V. Rakocevic, J. Habermann "Performance Comparison of Gateway Discovery Algorithms in Ad Hoc Networks with Mobile Nodes" Wireless Conference 2006, 12th European, April 2006.
- [10] H. Luo, R. Ramjee, P. Sinha, L. Li, S. Lu "UCAN: A Unified Cellular and AdHoc Network Architecture" MobiCom'03, San Diego, California, USA, September 2003.
- [11] A. Tahar, B. Azeddine. "AODV Extension Using Multi Point Relay for High Performance Routing in VANETS" International Journal of Computer Science Department, 2007.

การปรับปรุงการแบ่งโหลดเพื่อกระจายข้อมูลหลายทางบนเครือข่ายเฉพาะกิจเคลื่อนที่

Multipath Load-balancing Improvement on Mobile Ad-hoc Network

ศุภโชค พงษ์ขวัญ (Supachoke Pongquan) วรวัชร ณรงค์ชวณะ (Worrawat Narongkhachavana)

และ สุเมธ ประภาวัต (Sumet Prabhavat)

คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง กรุงเทพฯ

aquarix38425@gmail.com, worrawat@gmail.com, sumet@it.kmitl.ac.th

บทคัดย่อ

ความล้มเหลวของการสื่อสารบนเครือข่ายเฉพาะกิจ มักมีสาเหตุมาจากเส้นทางที่ใช้งานอยู่ถูกตัดขาด อันเนื่องมาจากการสูญหายของโหนดเคลื่อนที่ เช่น การเคลื่อนที่ของโหนดเคลื่อนที่ พลังงานที่จำกัดของโหนดเคลื่อนที่ เป็นต้น งานวิจัยนี้จึงพัฒนารูปแบบวิธีการส่งข้อมูล โดยใช้การส่งข้อมูลหลายเส้นทาง และวิธีการเลือกเส้นทางของโพรโทคอลซึ่งส่งผลให้สามารถใช้งานเส้นทางที่ถูกเลือกได้อย่างมีประสิทธิภาพ จากผลการจำลองการส่งข้อมูลระหว่างโหนดเคลื่อนที่ แสดงให้เห็นถึงการเพิ่มขึ้นของอัตราสำเร็จในการส่งข้อมูล และเวลาที่ใช้ในการทำงานลดลง

คำสำคัญ: เครือข่ายสื่อสารเฉพาะกิจ การส่งข้อมูลหลายเส้นทาง การกระจายโหลดของเครือข่าย

Abstract

Mobile Ad-hoc network has been popularly focused in recent years. Major problem of Mobile Ad-hoc network is the mobility of mobile nodes which could bring packet delivery failure into the network. This paper aims to develop multipath load-balancing algorithm into multipath routing protocol based on Dynamic Source Routing Protocol (DSR). The algorithm use multiple routes found from route discovery process and utilize that routes by using per-packet allocation scheme. We use extensive simulation to show the result which increase success ratio and reduce overall worktime.

Keyword: Mobile Ad-hoc Network, Dynamic Source Routing Load Balancing, Multipath Routing.

1. บทนำ

เทคโนโลยีการสื่อสารคือหัวใจสำคัญของระบบเครือข่าย ซึ่งมีทั้งระบบเครือข่ายแบบ Infrastructure ที่จำเป็นต้องพึ่งพาอุปกรณ์ที่ทำหน้าที่รับ-ส่ง ข้อมูล โดยหากเกิดเหตุไม่คาดฝันที่ทำให้ระบบ Infrastructure ไม่สามารถทำงานได้ ก็จะทำให้การสื่อสารขัดข้อง เป็นที่มาของระบบเครือข่ายไร้สายเฉพาะกิจ (Ad Hoc Network) ซึ่งเป็นระบบเครือข่ายแบบที่ไม่จำเป็นต้องมีอุปกรณ์ที่ทำหน้าที่รับ-ส่ง ข้อมูล เป็นตัวกลาง (Infrastructureless) เพื่อที่จะให้การติดต่อสื่อสารเป็นไปได้ในเวลาที่เกิดเหตุขัดข้องจนระบบ Infrastructure ไม่สามารถทำงานได้

Mobile Ad Hoc Network (MANET) [1] เป็นระบบเครือข่ายแบบ Infrastructureless ที่ทำงานด้วยเทคโนโลยีเครือข่ายไร้สาย โดยเทคโนโลยีที่แพร่หลายที่สุดก็คือเทคโนโลยี IEEE 802.11 หรือ Wi-Fi ซึ่งมีอยู่ในสมาร์ตโฟนทั่วไป ดังนั้นการนำเทคโนโลยี Wi-Fi มาใช้สำหรับสร้างเครือข่ายเฉพาะกิจของสมาร์ตโฟนจึงเป็นทางเลือกที่ดีในการทำให้เกิดเครือข่ายสื่อสารในสถานการณ์วิกฤตดังกล่าวได้

แต่ปัญหาของการสร้างเครือข่าย MANET ด้วยเทคโนโลยี Wi-Fi ก็คือ โหนดเคลื่อนที่ในเครือข่าย MANET มีโอกาสสูงที่จะขาดการเชื่อมต่อจากกัน เพราะเทคโนโลยี Wi-Fi ไม่เหมาะสมกับการใช้งานกับเครือข่ายที่โหนดเคลื่อนที่มี การเคลื่อนที่ เนื่องจากเทคโนโลยีดังกล่าวมีระยะเวลาเชื่อมต่อค่อนข้างสั้น ทำให้มีโอกาสที่โหนดเคลื่อนที่จะเคลื่อนที่ออกนอกจากระยะการเชื่อมต่อ ส่งผลให้มีโอกาสเกิดการขาดการเชื่อมต่อระหว่างโหนดเคลื่อนที่ต้นทางและปลายทาง ซึ่งนำไปสู่ปัญหาความยั่งยืนทนทานของเครือข่าย ดังนั้นการมีการ

หาเส้นทางหลายเส้นในเครือข่าย MANET สามารถลดโอกาสเกิดการขาดการเชื่อมต่อระหว่างโหนดต้นทางและปลายทางได้ จึงเป็นทางเลือกที่เป็นไปได้ในการแก้ปัญหาดังกล่าว

ในส่วนที่ 2 จะกล่าวถึงทฤษฎีพื้นฐานที่เกี่ยวข้องและนำมาพัฒนา, ส่วนที่ 3 มีการอธิบายรายละเอียดอัลกอริทึมที่นำมาใช้พัฒนาระบบเครือข่าย, ส่วนที่ 4 เป็นการประเมินผลการทดลองของอัลกอริทึมที่พัฒนาขึ้นมา โดยเปรียบเทียบกับโปรโตคอลหลักที่นำมาพัฒนา และส่วนที่ 5 เป็นการสรุปผลการทดลองของงานวิจัยนี้

2. วิจารณ์วรรณกรรม

2.1 Mobile Ad-hoc Network

ระบบเครือข่ายเฉพาะกิจไร้สาย (MANET) เป็นระบบที่ถูกออกแบบมาเพื่อรองรับการเชื่อมต่อระหว่างโหนดโดยไม่จำเป็นต้องพึ่งพาอุปกรณ์เชื่อมต่อพื้นฐาน เช่น แอคเซสพอยท์ สวิตช์ หรือ เราท์เตอร์ โดยตัวอุปกรณ์ (End-device) สามารถสร้างเครือข่ายขึ้นมาได้เอง และทำการติดต่อกันระหว่างโหนดด้วยตัวเอง เพื่อให้ข้อมูลส่งไปถึงปลายทาง โดยที่จุดเด่นของ MANET คือ สามารถจัดการกับอุปกรณ์ต่างๆ ที่มีการเคลื่อนที่ได้อย่างอิสระ (Node Mobility)

โปรโตคอลการค้นหาเส้นทางของระบบเครือข่ายเฉพาะกิจไร้สาย หากจัดแบ่งตามลักษณะการทำงานของโปรโตคอล [2], [3] จะแบ่งเป็น 2 กลุ่ม ได้แก่ 1) โปรโตคอลค้นหาเส้นทางแบบโปรแอกทีฟ ซึ่งในโปรโตคอลประเภทนี้โหนดทุกๆ ตัวในเครือข่ายจะมีตารางข้อมูลของโหนดต่างๆ (Routing Information) ภายในระบบ และมีการอัปเดตข้อมูลนั้นทุกๆ ช่วงเวลาที่กำหนด 2) โปรโตคอลค้นหาเส้นทางแบบรีแอคทีฟ จะไม่มีการกำหนดเวลาในการอัปเดตข้อมูลภายในช่วงเวลาที่กำหนด แต่จะทำการหาเส้นทางของทั้งระบบเฉพาะเวลาที่จะมีการส่งข้อมูลระหว่างโหนดตั้งแต่ต้นทางจนถึงปลายทาง

2.2 Dynamic Source Routing Protocol (DSR)

DSR [4] เป็นโปรโตคอลในการหาเส้นทางสำหรับการส่งข้อมูลแบบรีแอคทีฟ โดยเมื่อโหนดต้นทางที่ต้องการที่จะส่งข้อมูล โหนดต้นทางจะมีการ broadcast แพ็คเก็ตพิเศษที่เพิ่ม DSR Header เข้าไปในแพ็คเก็ตแบบปกติ เพื่อทำการเก็บข้อมูลของโหนดต่างๆ ภายในเครือข่าย ซึ่งตัวแพ็คเก็ตที่มี DSR Header จะเก็บข้อมูลไอพีแอดเดรส ของโหนดที่ส่งผ่านเข้าไปในส่วนหัวของแพ็คเก็ต และเมื่อพบโหนดปลายทางแล้วก็จะส่งข้อความตอบกลับมายังต้นทาง จากนั้นก็จะเริ่มทำการสร้างเส้นทางเพื่อทำการส่งแพ็คเก็ตปกติตามเส้นทางที่ได้ทำการบันทึกไว้

โปรโตคอล DSR มีกลไกการทำงานหลักสองกลไก คือ 1) กลไกการค้นหาเส้นทาง (Route Discovery) จะเริ่มต้นทำงานเมื่อโหนดต้นทางต้องการส่งแพ็คเก็ตไปยังปลายทางแต่ไม่มีข้อมูลเส้นทางที่จะทำการส่งแพ็คเก็ตนั้นๆ ไปยังปลายทาง จึงเริ่มทำการค้นหาเส้นทาง โดยโหนดต้นทางจะทำการ broadcast แพ็คเก็ตร้องขอเส้นทาง (Route Request: RREQ) ออกไป โดยมีหมายเลข RREQ ID กำหนดไว้และเมื่อโหนดภายในเครือข่ายได้รับก็จะทำการตรวจสอบแพ็คเก็ตว่ามีปลายทางเป็นของตนเองหรือไม่ หากไม่ใช่ก็จะทำการ broadcast แพ็คเก็ตนั้นต่อออกไป และหากมีปลายทางเป็นตนเองก็จะทำการส่งแพ็คเก็ตตอบรับ (Route Reply: RREP) กลับมายังโหนดต้นทาง 2) กลไกการปรับปรุงเส้นทาง (Route Maintenance) จะถูกเรียกใช้เมื่อโหนดภายในเครือข่ายทำการส่งข้อมูลอยู่และเส้นทางที่ถูกใช้มีโหนดขาดหายไปจากเส้นทาง ทำให้การส่งข้อมูลหยุดชะงักลง โหนดระหว่างทางที่เชื่อมต่อกับโหนดที่หายไปจะทำการส่งแพ็คเก็ตแจ้งกลับมายังต้นทาง (Route Error: RERR) ว่าเส้นทางที่ใช้อยู่ได้ขาดลงไปแล้ว โหนดต้นทางก็จะทำการเรียกใช้เส้นทางอื่นๆ ที่อยู่ภายในแคชของตน หรืออาจเรียกกลไก Route Discovery เพื่อทำการหาเส้นทางใหม่

2.3 Multiple Source Routing Protocol (MSR)

โปรโตคอล MSR [5] พัฒนาขึ้นโดยใช้หลักการของ Dynamic Source Routing Protocol และนำหลักการของการส่งแบบหลายเส้นทางมาประยุกต์ใช้ได้ง่าย เนื่องจากในการทำงานปกติของ DSR เป็นแบบ Source Routing ซึ่งมีกลไกในการหา

เส้นทางก่อนส่งข้อมูลและมีการเก็บข้อมูลเส้นทางที่หามาได้ไว้กับโมบายล์โหนดที่ทำการส่งข้อมูล ทำให้การประมวลผลการส่งแบบหลายเส้นทางมีความเป็นไปได้ที่จะนำมาใช้กับโปรโตคอล DSR

กลไกการค้นหาเส้นทางในโปรโตคอล MSR มีการทำงานเหมือนในโปรโตคอล DSR แต่จะมีการนำข้อมูลเส้นทางที่ได้รับกลับมาจากโหนดปลายทางมาเก็บไว้ในแคชของตน และนำข้อมูลเส้นทางต่างๆ ที่ได้รับออกมาใช้ในการส่งแพ็คเก็ตโดยใช้รูปแบบการส่งแพ็คเก็ตแบบ Round Robin [6] ซึ่งจะมีการวนใช้งานแต่ละเส้นทางไปเรื่อยๆ

3. ขั้นตอนการดำเนินงานวิจัย

จากการศึกษาทบทวนวรรณกรรมในส่วนที่ 2 เกี่ยวกับเครือข่ายเฉพาะกิจเคลื่อนที่, โปรโตคอล DSR และ โปรโตคอล MSR ผู้วิจัยได้สังเกตเห็นว่าโปรโตคอล DSR ที่มีการใช้เส้นทางเพียงเส้นทางเดียวที่ได้รับจากกระบวนการ Route Discovery นั้นมีโอกาสทำให้เกิดปัญหา Single-point-of-failure ซึ่งมีผลต่อความน่าเชื่อถือของโปรโตคอล โดยชีวิตได้จากอัตราสำเร็จในการส่งแพ็คเก็ต นอกจากนี้การเลือกใช้เส้นทางเพียงเส้นทางเดียว ยังมีโอกาสทำให้เกิดความคับคั่งภายในเส้นทางนั้น ในขณะที่ยังมีเส้นทางอื่นที่ได้รับมาสามารถใช้ในการส่งข้อมูลได้อีก ส่งผลให้ใช้เวลานานในการส่งข้อมูลสำเร็จ

ผู้วิจัยจึงทำการนำเสนอแนวคิดการพัฒนาโปรโตคอล MSR ที่มีอยู่เดิมโดยจะปรับปรุงรูปแบบการส่งแพ็คเก็ตเดิมที่ใช้แบบ Round Robin ผู้วิจัยนำเสนอรูปแบบการส่งแพ็คเก็ตแบบ Weight-Round Robin (WRR) แทนการใช้รูปแบบการส่งแบบเดิม เนื่องจากการส่งแพ็คเก็ตแบบ Round Robin นั้นมีการกระจายโหลดของเส้นทางเท่าๆ กัน จึงมีโอกาสที่เส้นทางที่มีดีเลย์สูงจะถูกเรียกใช้งานเท่ากับเส้นทางที่มีดีเลย์ต่ำ ซึ่งค่าดีเลย์วัดได้จากกการจับเวลาตั้งแต่ส่งแพ็คเก็ต Route Request ออกไปและได้รับ Route Reply กลับมาผ่านเส้นทางนั้นๆ

ผู้วิจัยได้ทำการนำค่าบิตเรทจากค่าดีเลย์ของเส้นทางขึ้นมาเป็นตัวแปรในการกำหนดน้ำหนักให้กับเส้นทางบนเน็ตเวิร์คโดยทั้งหมดนี้ ผู้วิจัยคาดว่าค่าบิตเรทดังกล่าวสามารถประเมินคุณภาพของเส้นทางได้โดยประมาณ และมีการใช้จำนวนฮอปของเส้นทางเป็นตัวกำหนดน้ำหนักของในกรณีที่มีค่าบิตเรทของเส้นทางเท่ากันดังที่แสดงให้เห็จากโค้ดเทียมในภาพที่ 1

```

1 struct route
2 {
3     path
4     bitrate
5     num_hop
6 };
7 new vector srt_vector
8 new vector weight_vector
9 if (srt_vector.size() = 0)
10 {
11     srt_vector.pushback(route);
12 }
13 else
14 {
15     for(int i=0; i<srt_vector.size();i++)
16     {
17         if (srt_vector.at(i).bitrate < route.getBitrate()
18             OR srt_vector.at(i).num_hop < route.num_hop)
19         {
20             srt_vector.insert(srt_vector.at(i-1), route);
21             weight_vector.insert(weight_vector.at(i-1), i);
22         }
23     }
24     if(n==srt_vector.size())
25     {
26         srt_vector.pushback(route);
27         weight_vector.pushback(srt_vector.size());
28     }
29 return srt_vector;

```

ภาพที่ 1: แสดงโค้ดเทียมของกระบวนการหาเส้นทาง การกำหนดค่าตัวแปรต่างๆ ดังนี้

$route$ = โครงสร้างของเส้นทาง ประกอบไปด้วย

- $path$ = เส้นทาง
- $bitrate$ = บิตเรทของค่าดีเลย์
- num_hop = จำนวนฮอป

srt_vector = เวกเตอร์ที่เก็บเส้นทางต่างๆ ไว้

$weight_vector$ = เวกเตอร์ที่เก็บน้ำหนักของเส้นทาง

n = ขนาดของเวกเตอร์

การทำงานของโค้ดข้างต้นจะเริ่มทำงานเมื่อเริ่มกระบวนการค้นหาเส้นทางโดย กำหนดค่าจำนวนเต็ม n ให้เท่ากับขนาดของเวกเตอร์เส้นทาง เมื่อเริ่มต้นค้นหาเส้นทางหากพบเส้นทางที่ใช้ได้ก็จะทำการเพิ่ม $route$ เข้าไปในเวกเตอร์ และหากพบเส้นทางต่อมาก็จะทำการเปรียบเทียบค่าบิตเรท ($bitrate$) และจำนวนฮอป (num_hop) ของเส้นทางภายในเวกเตอร์ โดยจะกำหนดให้เส้นทางที่มีค่าบิตเรทสูงกว่าอยู่ด้านหน้าของเวกเตอร์ และหากมีค่าบิตเรทเท่ากันก็จะทำการเปรียบเทียบจำนวนฮอป โดยให้เส้นทางที่ใช้จำนวนฮอปน้อยกว่าอยู่ด้านหน้า พร้อมกับเพิ่มน้ำหนักให้กับเส้นทางเข้าไปยัง $weight_vector$ ที่ตำแหน่งเดียวกับเส้นทางนั้นๆ เมื่อทำงานเสร็จสิ้นก็จะทำการส่งค่า srt_vector ที่เก็บเส้นทางต่างๆ ออกไปใช้งานต่อ

ในงานวิจัยนี้ ผู้วิจัยทำการคำนวณน้ำหนักของเส้นทางโดยกำหนดน้ำหนักสูงสุดตามจำนวนเส้นทางที่สามารถใช้ได้ลดลง

มาจนถึงเส้นทางที่มีน้ำหนักราคาต่ำสุดคือ 1 และมีการกระจายโหนด โดยใช้การกระจายตามแต่ละแพ็คเก็ต โดยจะส่งแพ็คเก็ตออกไปยังเส้นทางที่มีน้ำหนักสูงสุดเป็นจำนวนตามค่าน้ำหนักของเส้นทางนั้นๆ จากนั้นจึงส่งแพ็คเก็ตออกไปยังเส้นทางถัดไปตามค่าน้ำหนักของเส้นทางเช่นกัน

ผู้วิจัยสังเกตเห็นว่าการส่งแพ็คเก็ตแบบ WRR สามารถแก้ปัญหาดังกล่าวได้โดยใช้ค่าเฉลี่ยของเส้นทางเป็นตัวกำหนด อัตราการถ่วงน้ำหนักของการส่งแพ็คเก็ตทำให้สามารถแบ่งโหนดของการส่งแพ็คเก็ตไปยังเส้นทางที่มีค่าน้ำหนักน้อยกว่าได้ดี ทั้งนี้เพื่อลดเวลาการทำงานโดยรวมของระบบตั้งแต่เริ่มส่งแพ็คเก็ตจนกระทั่งส่งแพ็คเก็ตทั้งหมดสำเร็จ โดยจะมีการนำเสนอผลการทดลองในส่วนถัดไป

4. ผลการทดลอง

จากการนำเสนอการปรับปรุงโปรโตคอลในส่วนที่ 3 ผู้วิจัยได้ทำการทดลองโดยใช้โปรแกรมจำลองระบบเครือข่าย และนำผลการจำลองเครือข่ายมาเปรียบเทียบกับอัลกอริทึม DSR และ MSR โดยมีข้อกำหนดควบคุม และผลการทดลองดังนี้

4.1 การควบคุมข้อกำหนดระบบเครือข่ายของแต่ละโปรโตคอล

การทดลองมีการกำหนดการควบคุมโดยมีการกำหนดให้โมบายล์โหนดมีการเคลื่อนที่อยู่เสมอด้วยความเร็วคงที่ โดยที่ความเร็วของแต่ละโหนดมีการสุ่มแบบยูนิฟอร์ม มีการกระจายตัวของโหนดแบบสุ่มภายในพื้นที่กว้าง 1500 เมตร ยาว 1000 เมตร กำหนดให้โหนดต้นทางส่งข้อมูลไปยังโหนดปลายทางด้วยโปรโตคอล UDP และให้ระบบหยุดทำงานเมื่อโหนดปลายทางได้รับข้อมูลจำนวน 1000 แพ็คเก็ต โดยมีการแสดงข้อมูลจำเพาะอื่นๆ ภายในตารางที่ 1 ต่อไปนี้

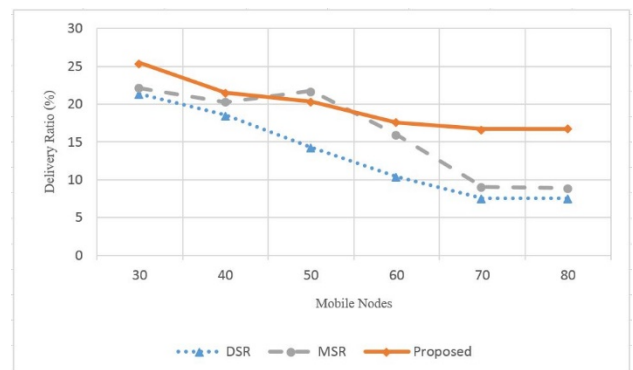
ตารางที่ 1: แสดงข้อมูลจำเพาะของการจำลองการทำงาน

พื้นที่จำลอง	1500 เมตร × 1000 เมตร
รัศมีการทำงานของโหนด	250 เมตร
จำนวนแพ็คเก็ตที่ปลายทางได้รับ	1000 แพ็คเก็ต
จำนวนโหนดในเครือข่าย	30, 40, 50, 60, 70, 80
อัตราการส่งแพ็คเก็ต	4 แพ็คเก็ต/วินาที
โปรโตคอลที่ใช้ส่งข้อมูล	UDP
ความเร็วในการเคลื่อนที่เฉลี่ย	2.5, 5, 7.5, 10 เมตร/วินาที
จำนวนรอบที่ใช้ในการทดลอง (ต่อหนึ่งรูปแบบจำลอง)	40 รอบ

4.2 ตัวชี้วัดและผลการทดลอง

จากการควบคุมข้อกำหนดต่างๆ ในส่วนที่ 4.1 และทำการทดลองจำลองการทำงานของอัลกอริทึมการค้นหาเส้นทางของโปรโตคอล DSR โปรโตคอล MSR และอัลกอริทึมที่ผู้วิจัยนำเสนอ ผู้วิจัยได้เสนอผลการจำลองระบบออกมาในรูปแบบกราฟโดยมีตัวชี้วัดต่างๆ ที่ใช้ในการเปรียบเทียบผลการทำงานคือ ผลสำเร็จของการส่งแพ็คเก็ต และเวลาการทำงานตั้งแต่เริ่มส่งแพ็คเก็ตจนโมบายล์โหนดปลายทางได้รับแพ็คเก็ตเป็นจำนวน 1,000 แพ็คเก็ต

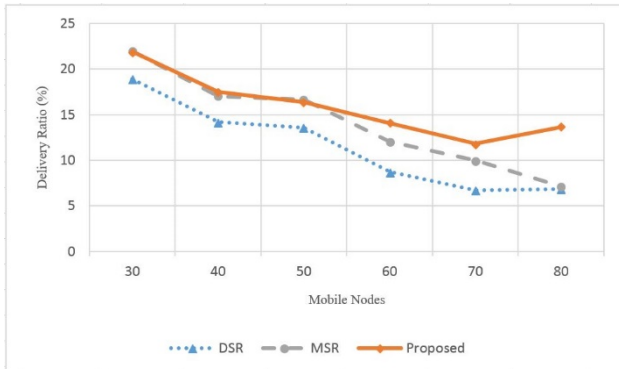
การทดลองควบคุมการทำงานโดยกำหนดความเร็วในการเคลื่อนที่เฉลี่ยเป็น 7.5 และ 10 เมตร/วินาที พบว่าอัตราสำเร็จในการส่งแพ็คเก็ตของระบบที่มีจำนวนโมบายล์โหนดที่เพิ่มขึ้นจะส่งผลให้อัตราสำเร็จในการส่งแพ็คเก็ตลดลง โดยในอัลกอริทึมของ MSR และอัลกอริทึมที่ผู้วิจัยนำเสนอ สามารถทำงานได้ดีกว่าอัลกอริทึมการค้นหาเส้นทางของ DSR ตามภาพที่ 1 และภาพที่ 2 ซึ่งจะเห็นได้ว่าเมื่อมีจำนวนโมบายล์โหนดเยอะ อัลกอริทึมที่นักวิจัยได้นำเสนอจะสามารถทำงานได้ดีกว่าอัลกอริทึม MSR และ DSR ที่ยังมีค่าใกล้เคียงกัน ทั้งนี้เนื่องจากกลไกการกระจายโหนดที่ทำให้สามารถใช้งานเส้นทางที่มีอยู่ได้อย่างมีประสิทธิภาพ



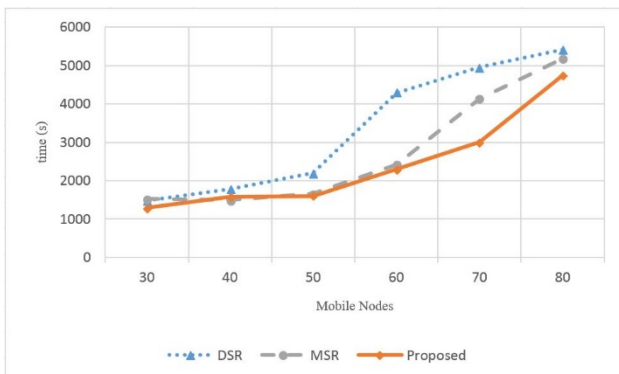
ภาพที่ 1: แสดงผลการทดลองอัตราสำเร็จต่อจำนวนโมบายล์โหนดที่ความเร็ว 7.5 เมตร/วินาที

ในการทดลองเพื่อหาเวลาการทำงานด้วยการกำหนดความเร็วในการเคลื่อนที่เฉลี่ยเป็น 7.5 และ 10 เมตร/วินาที ตามภาพที่ 3 และภาพที่ 4 แสดงให้เห็นว่าที่จำนวนโหนดตั้งแต่ 30 – 50 โหนด แต่ละอัลกอริทึมจะใช้เวลาในการทำงานใกล้เคียงกัน แต่เมื่อจำนวนโหนดเพิ่มขึ้น อัลกอริทึมที่นักวิจัยนำเสนอสามารถทำงานได้ดีกว่าโปรโตคอล MSR และ DSR อย่างเห็นได้ชัด ทั้งนี้เพราะว่ามีการกำหนดน้ำหนักให้กับเส้นทาง ทำให้

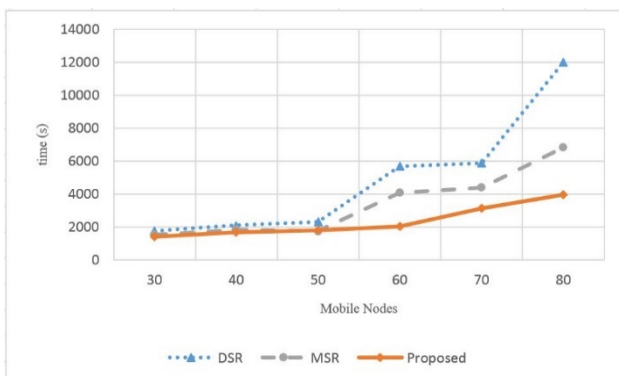
สามารถกระจายโหนดออกไปยังเส้นทางที่มีบิตเรตสูงกว่าสามารถระบายการไหลของแพ็คเก็ตได้และไม่เกิดความคับคั่งภายในเส้นทางที่มีบิตเรตต่ำ



ภาพที่ 2: แสดงผลการทดลองอัตราสำเร็จต่อจำนวนโหนดที่ความเร็ว 10 เมตร/วินาที



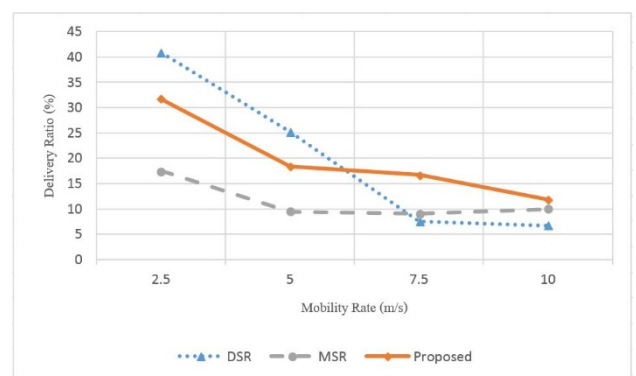
ภาพที่ 3: แสดงผลการทดลองเวลาการทำงานต่อจำนวนโหนดที่ความเร็ว 7.5 เมตร/วินาที



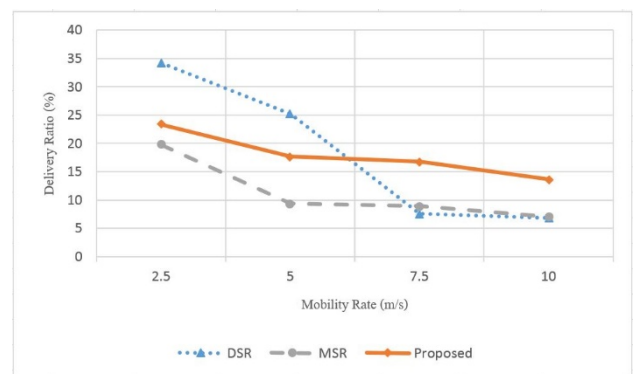
ภาพที่ 4: แสดงผลการทดลองเวลาการทำงานต่อจำนวนโหนดที่ความเร็ว 10 เมตร/วินาที

ภาพที่ 5 และภาพที่ 6 แสดงการเปรียบเทียบการทำงานที่มีจำนวนโหนดภายในเครือข่ายสูง เช่น 70 และ 80 โหนด ระหว่างความเร็วเฉลี่ยในการเคลื่อนที่ของโหนดที่ 7.5 และ 10 เมตร/วินาที อัตราสำเร็จในการส่งแพ็คเก็ตของโหนด จะเห็นได้ว่าเมื่อโหนดมีการเคลื่อนที่ด้วยความเร็วต่ำ ประสิทธิภาพการ

ทำงานของอัลกอริทึม DSR จะทำงานได้ดีกว่าอัลกอริทึม MSR และอัลกอริทึมที่แนะนำอื่น ๆ แต่เมื่อโหนดเคลื่อนที่ด้วยความเร็วสูงขึ้น อัตราสำเร็จในการส่งแพ็คเก็ตของ DSR จะลดลงอย่างมาก ในขณะที่อัลกอริทึมของผู้วิจัยมีการลดลงค่อนข้างน้อย ทำให้สามารถทำงานได้ดีกว่าอัลกอริทึม DSR และ MSR ที่ความเร็ว 7.5 และ 10 เมตร/วินาที ทั้งนี้เป็นผลมาจากการเคลื่อนที่ของโหนดด้วยความเร็วสูง ทำให้มีโอกาสสูงที่เส้นทางแพ็คเก็ตจะถูกตัดขาด ทำให้การแพ็คเก็ตในเครือข่ายที่ใช้อัลกอริทึม DSR หยุดชะงักลง ในขณะที่ MSR และอัลกอริทึมที่แนะนำยังสามารถส่งแพ็คเก็ตได้อยู่เนื่องจากการใช้เส้นทางส่งแพ็คเก็ตหลายเส้นทาง



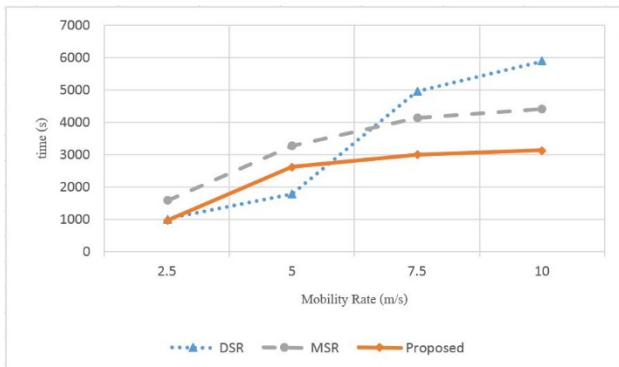
ภาพที่ 5: แสดงผลการทดลองอัตราสำเร็จต่อความเร็วเฉลี่ยสำหรับโหนดจำนวน 70 โหนด



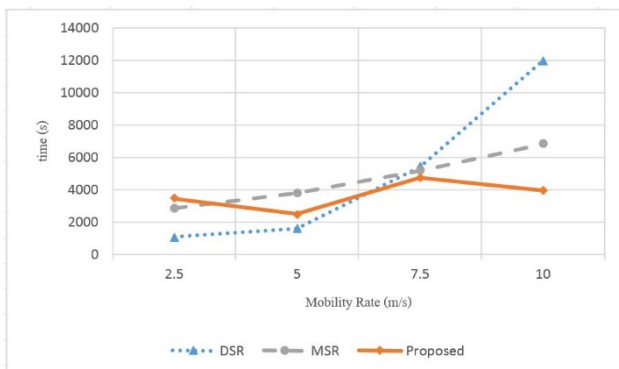
ภาพที่ 6: แสดงผลการทดลองอัตราสำเร็จต่อความเร็วเฉลี่ยสำหรับโหนดจำนวน 80 โหนด

ภาพที่ 7 และภาพที่ 8 เป็นการเปรียบเทียบเวลาการทำงานต่อความเร็วในการเคลื่อนที่ของโหนดของเครือข่ายที่มีโหนดจำนวน 70 และ 80 โหนดตามลำดับ อัลกอริทึมของ DSR สามารถทำงานได้ดีกว่าในความเร็ว 2.5 และ 5 เมตร/วินาที แต่เมื่อความเร็วเพิ่มขึ้นเป็น 7.5 และ 10 เมตร/วินาที อัลกอริทึม DSR จะใช้เวลาทำงานเพิ่มขึ้นอย่างมาก ส่วนอัลกอริทึม MSR และอัลกอริทึมที่แนะนำใช้เวลา

เพิ่มขึ้นน้อยกว่าอัลกอริทึม DSR เป็นผลมาจากการที่เส้นทางที่ส่งแพ็คเก็ตอยู่ถูกตัดขาด ทำให้อัลกอริทึม DSR ที่ทำงานโดยการส่งแพ็คเก็ตเกิดเส้นทางเดียว ต้องเริ่มกระบวนการค้นหาเส้นทางใหม่ จึงใช้เวลาเพิ่มขึ้นเป็นอย่างมาก ในขณะที่อัลกอริทึม MSR และอัลกอริทึมที่นำเสนอ ยังมีเส้นทางอื่นๆ ที่สามารถใช้ได้มารองรับการส่งแพ็คเก็ตไปยัง โมงบายล์ โหนดปลายทาง



ภาพที่ 7: แสดงผลการทดลองเวลาการทำงานต่อความเร็วเฉลี่ยของ โมงบายล์ โหนดจำนวน 70 โหนด



ภาพที่ 8: แสดงผลการทดลองเวลาการทำงานต่อความเร็วเฉลี่ยของ โมงบายล์ โหนดจำนวน 80 โหนด

5. สรุปผลการทำงานของระบบ

นักวิจัยได้นำเสนอการพัฒนาอัลกอริทึมในการค้นหาเส้นทางของเครือข่ายเฉพาะกิจเคลื่อนที่ โดยมีโปรโตคอล DSR เป็นหลัก และนำโปรโตคอล MSR ที่พัฒนาต่อมาจาก DSR มาเป็นพื้นฐานในการพัฒนาอัลกอริทึม ซึ่งได้ทำการเพิ่มอัลกอริทึมการค้นหาเส้นทางโดยใช้ค่าบิเทรทจากดาต้าลิงค์เลเยอร์ และจำนวนฮอปของเส้นทางมาเป็นตัวกำหนดน้ำหนักให้กับเส้นทาง

จากผลการทดลองในส่วนที่ 4 สามารถบอกได้ว่าอัลกอริทึมที่นักวิจัยนำเสนอ นั้น สามารถทำงานได้ดีกว่าอัลกอริทึม DSR

และ MSR ในสถานการณ์ที่มีจำนวน โมงบายล์ โหนดสูง หรือมีการเคลื่อนที่ด้วยความเร็วมาก เป็นผลมาจากการเพิ่มอัลกอริทึมค้นหาเส้นทางให้มีการกำหนดน้ำหนักให้กับเส้นทาง ทำให้ไม่เกิดความคับคั่งในเส้นทางใดเส้นทางหนึ่ง และสามารถดึงประสิทธิภาพของ โมงบายล์ โหนดออกมาใช้งานได้อย่างเต็มที่

เอกสารอ้างอิง

- [1] T. S. Rappaport. "Wireless Communications Principles & Practice." Upper Saddle River, New Jersey: Prentice-Hall, Inc.1996
- [2] T. Lin. "Mobile Ad-hoc Network Routing Protocols: Methodologies and Applications." Dissertation submitted to the Faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the Ph.D. in Computer Engineering. 2004
- [3] K.Gorantala. "Routing Protocols in Mobile Ad-hoc Networks" Master's Thesis in Computing Science at Umeªa University. June 15, 2006
- [4] D. B. Johnson and D. A. Maltz, "Dynamic Source Routing in Ad Hoc Wireless Networks." Computer Science Department, Carnegie Mellon University. 1996
- [5] L. Wang, Y. Shu, M. Dong and L. Zhang, "Adaptive Multipath Source Routing in Ad Hoc Networks." Department of Computer Science, Tianjin University, Tianjin, China
- [6] S. Prabhavat, H. Nishiyama, N. Ansari and N. Kato "On Load Distribution over Multipath Networks." *IEEE Communications Survey & Tutorials*, vol. 14, no. 3, pp. 662-680, 2012.

การป้อนข้อมูลย้อนกลับขนาดต่ำเพื่อปรับชดเชยสัญลักษณ์ล่วงหน้าสำหรับช่องสัญญาณ มีการจางหายแบบพหุวิถี ในระบบโอเอฟดีเอ็ม Low information feedback and Makeup constellation method OFDM system over Multipath fading

ชัยพร เขมะภตะพันธ์(Chaiyaporn Khemapatapan)¹ และ รัชพล เพ็ญโบราณ(Rushapon Piaboran)²

สาขาวิศวกรรมคอมพิวเตอร์และโทรคมนาคม คณะวิศวกรรมศาสตร์ มหาวิทยาลัยธุรกิจบัณฑิตย์

¹chaiyaporn@dpu.ac.th, ²rushapon@hotmail.com

บทคัดย่อ

บทความนี้ศึกษาและพัฒนาระบบการสื่อสารแบบ ไร้สาย ระบบโอเอฟดีเอ็มสำหรับช่องสัญญาณแบบพหุวิถี ซึ่งมีการประยุกต์การปรับชดเชยสัญลักษณ์ทั้งขนาดและเฟสของสัญญาณ ล่วงหน้า โดยใช้บิตนำร่องเพื่อทดสอบคุณลักษณะของช่องสัญญาณ จากนั้นภาครับจะส่งข้อมูลย้อนกลับที่มีอัตราต่ำมากกลับมายังภาคส่งเพื่อใช้ปรับสัญลักษณ์ล่วงหน้าโดยทำการชดเชยและจัดกลุ่มสัญลักษณ์ใหม่ก่อนส่งสัญญาณ สัญลักษณ์ดังกล่าวออกไปในช่องสัญญาณที่เป็นแบบพหุวิถี ผลลัพธ์ที่ได้จากการจำลองระบบที่มีการปรับชดเชยร่วมกับอควไลซ์ที่เครื่องรับเปรียบเทียบกับการทำงานที่มีอควไลเซอร์เพียงอย่างเดียว พบว่าระบบที่เสนอสามารถลดอัตราบิตที่ผิดพลาดได้เป็นอย่างดีและให้ผลการทำงานดีกว่าระบบ THP เล็กน้อย

คำสำคัญ โอเอฟดีเอ็ม การเข้ารหัสล่วงหน้า การชดเชย

Abstract

This paper studied and improved wireless OFDM communication system over multipath channel. Makeup constellation method is applied both amplitude and phase of transmitted signal to improve the performance. Pilot information bits are used to trace the channel variation. The receiver sends back the transmitter the low rate information in order to use for making up a constellation before transmission. The results from the simulations

based on 2 scenarios: makeup constellation with equalizer and only equalizer found that the proposed system can effectively reduce bit error rate and the performance is slightly less than THP's performance.

Keyword: OFDM System, Pre-coding, Make up, Multipath Fading channel.

1. บทนำ

ปัจจุบันความต้องการในการเพิ่มความเร็วยุติการรับส่งข้อมูลนั้น เป็นที่สนใจอย่างต่อเนื่อง ความมุ่งหวังในการสร้างอุปกรณ์ที่มีความสามารถส่งผ่านข้อมูลในระยะไกลๆและมีอัตราข้อผิดพลาดต่ำด้วยขบวนการที่ไม่ซับซ้อน เป็นที่สนใจเป็นอย่างมาก ซึ่งในระบบการสื่อสารไร้สายนั้น ได้มีการพัฒนาระบบการส่งสัญญาณแบบ OFDM (Orthogonal Frequency Division Multiplex) ซึ่งเป็นลักษณะ MCM (Multiple Carrier Modulation) [1] ที่มีการเพิ่มความหนาแน่นของสัญญาณพาหะย่อยและการซ้อนทับสัญญาณพาหะที่อยู่ติดกัน โดยเทคนิคการมอดูเลชันแบบหลายคลื่นพาหะ และเป็นรูปแบบของการสื่อสารแบบขนาน

OFDM เป็นเทคโนโลยีที่ใช้สำหรับถ่ายทอด สัญญาณผ่านทางคลื่นวิทยุ ซึ่งกระบวนการนี้จะเป็นการแบ่งสัญญาณ ความถี่ออกเป็นหลาย ๆ เซ็กเมนต์ และในแต่ละเซ็กเมนต์จะมีความถี่เฉพาะตัว ซึ่งจะทำให้สามารถรองรับความจุและอัตราการส่งข้อมูลที่สูงกว่าและใช้แบนวิดท์ได้เต็มประสิทธิภาพ แต่การส่งสัญญาณจากต้นทางไปยังปลายทาง โดยตัวกลางที่เป็นอากาศ

เพื่อนำพาข้อมูลข่าวสาร (Information) เมื่อสัญญาณเดินทางผ่านตัวกลาง ก็จะมีการลดทอน (Attenuation) ของสัญญาณเกิดขึ้น โดย “ขนาด” ของการลดทอนก็จะขึ้นอยู่กับสถานะของช่องสัญญาณนั้นๆ และสัญญาณ

การเดินทางจากต้นทางไปยังปลายทางของสัญญาณ จะไม่ได้มีเพียงแค่คลื่นที่เป็นเส้นตรงเพียงอย่างเดียว แต่กลับมีการสะท้อนไปมาระหว่าง ดึกสูงหรือสิ่งก่อสร้าง หรือสภาพแวดล้อมอื่นๆ ที่สามารถสะท้อนได้ เช่น ยานพาหนะ เป็นต้น ทำให้สัญญาณที่มาถึงเครื่องรับประกอบด้วยสัญญาณสะท้อนมากกว่าหนึ่งสัญญาณ โดย ค่าขนาด และ ค่าหน่วยเวลาของแต่ละสัญญาณ ก็จะมีค่าแตกต่างกันไป โดยสัญญาณอันเกิดจากการสะท้อนหลายๆสัญญาณนั้นจะเรียกรวมๆว่าเป็นพหุวิถี หรือ “มัลติพาท (Multipath)” ดังที่ได้กล่าวมานี้เป็นผลให้สัญญาณที่ปลายทางผิดเพี้ยนไป จึงนำได้มีหาวิธีการต่างๆ ที่สามารถลดความผิดเพี้ยนของข้อมูลปลายทางตามสภาพแวดล้อมต่างๆ โดยหลักการ ทำงานจะใช้ การจำแนกประเภทของวงจรปรับเท่าหรืออิกวอลไลเซอร์ (Equalizer) [2] จากนั้นก็นำค่าของ ISI (Inter symbol interference) ที่ประมาณได้นี้ ไปหักลบกับสัญญาณข้อมูลที่ตามมาภายหลังเพื่อกำจัดค่า ISI ก่อนที่จะทำการตรวจจับสัญญาณและทำการตัดสินใจ แต่สำหรับช่องการสื่อสารโทรศัพท์เคลื่อนที่มีปัญหาความผิดเพี้ยนสูงมาก ทำให้ไม่สามารถนำวงจรมาใช้ในทางปฏิบัติ ด้วยเหตุนี้จึงมีการพัฒนาวิธีที่ดีกว่า ซึ่งต่างจากวิธีเดิมคือกระบวนการดึงสัญญาณกลับมาจะมีการพิจารณาสัญญาณครวละมาๆ แทนการพิจารณาสัญญาณทีละค่า โดยวิธีการต่อเนื่องนี้ทุกรูปแบบที่เป็นไปได้เพื่อหาชุดสัญญาณที่ดีที่สุด ดังนั้นสัญญาณในลักษณะนี้ ต้องอาศัยการคำนวณเป็นจำนวนมาก

บทความนี้จะขอนำเสนอการป้อนข้อมูลย้อนกลับขนาดต่ำเพื่อปรับชดเชยสัญญาณสำหรับช่องสัญญาณแบบพหุวิถีในระบบการส่งสัญญาณแบบ โอเอฟดีเอ็ม ที่มีการจางหายแบบพหุวิถี โดยมีการป้อนข้อมูลย้อนกลับในอัตราที่ต่ำ (Low

Information Feedback) นำจัดกลุ่มสัญญาณแบบปรับตัวได้โดยใช้บิตย้อนกลับมายังต้นทางเพื่อปรับกลุ่มสัญญาณใหม่ พร้อมทั้งการปรับขนาดของเฟสและแอมพลิจูดของสัญญาณที่จะส่งไปยังปลายทางทำให้ได้สัญญาณที่ดีขึ้น โดยใช้แบบจำลองเบสแบนด์ ในส่วนขององค์ประกอบด้านเครื่องรับสัญญาณนั้นจะมีตัวปรับชดเชยสัญญาณล่วงหน้า พร้อมด้วยวงจรปรับเท่าหรืออิกวอลไลเซอร์ (Equalizer) แบบ Zero Forcing เพื่อเทียบกับแบบจำลองเบสแบนด์ ที่ไม่มีปรับกลุ่มสัญญาณใหม่ เพื่อให้ทราบถึงข้อแตกต่างของอัตราบิดพลิพลาด ข้อดีของระบบนี้ก็คือช่วยลดการใช้แบนด์วิดท์ในการป้อนข้อมูลย้อนกลับ ลดอัตราการบิดพลิพลาดของข้อมูลในภาครับเป็นระบบที่ไม่มีควมซับซ้อนเข้าใจได้ง่าย และมีการใช้งานแบนด์วิดท์อย่างเต็มประสิทธิภาพ

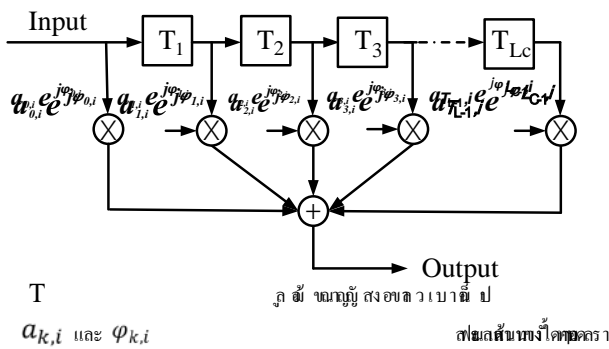
2. งานวิจัยที่เกี่ยวข้อง

สิ่งที่เกิดในระบบการสื่อสารแบบไร้สายนั้นคือ ช่องการสื่อสารที่มีการเปลี่ยนแปลงอยู่ตลอดเวลาอันเนื่องมาจากองค์ประกอบต่างๆ เช่น สิ่งแวดล้อม ความเร็วในการเคลื่อนที่ของเครื่องรับ-ส่ง รวมถึงอุณหภูมิ มีผลทำให้อุปกรณ์ตรวจจับสัญญาณ ได้รับข้อมูลที่ผิดพลาดอีกด้วย ซึ่งในปัจจุบันได้มีการพัฒนาอุปกรณ์ตรวจชนิดต่างเพื่อแก้ไขในจุดนี้อย่างกว้างขวาง [3],[4] ซึ่งในงานวิจัยชิ้นนี้ได้เป็นอีกแนวทางหนึ่งซึ่งเป็นการผสมผสานการทำงานระหว่างระบบการปรับแต่งชดเชยสัญญาณล่วงหน้าที่ใช้ร่วมกันกับระบบ OFDM ที่นิยมใช้กันในกรณีที่มีการพิจารณาถึง ลักษณะของความผิดเพี้ยนของสัญญาณที่รับได้ ณ ปลายทางเนื่องจากการถูกรบกวนจากสัญญาณรบกวนในระบบสื่อสารนั้น ในทางคณิตศาสตร์สามารถแสดงลักษณะของสัญญาณรบกวนที่เกิดขึ้นในระบบสื่อสารได้ โดยใช้ตัวแปรสุ่มแบบเกาส์ โดยที่เมื่อพิจารณาถึงสัญญาณข้อมูลที่รับได้ ณ ปลายทางนั้น จะมีค่าเท่ากับ

$$\begin{aligned} \bar{R}_i &= h(t)\bar{S}_i + \bar{n} \\ &= \sum_{k=0}^{L_c-1} (a_{k,i} e^{j\varphi_{k,i}} \bar{S}_k) + \bar{n}_i \end{aligned} \quad (1)$$

โดยที่ $a_{k,i}$ และ $\varphi_{k,i}$ เป็นตัวแปรที่แสดงถึงขนาดและเฟสที่เกิดการจางหายของสัญลักษณ์ของช่องสัญญาณแบบพหุวิธี L_c เป็นจำนวนพหุวิธีของสัญญาณที่ถูกนำมาคำนวณ

จากสมการที่ (1) จึงสามารถนำไปใช้ในการสร้างแบบจำลองช่องสัญญาณไร้สาย $h(t)$ คือช่องสัญญาณแบบพหุวิธีที่มีคุณสมบัติตามรูปที่ 1 โดย



รูปที่ 1: แบบจำลองของช่องสัญญาณที่นำมาพิจารณา

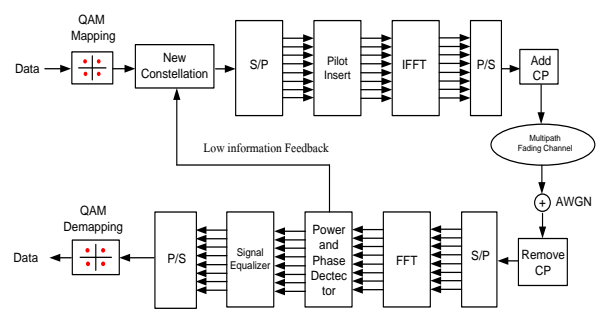
ในขณะที่ \bar{n} เป็นตัวแปรสุ่มแบบเกาส์ ที่มีลักษณะของ Probability Density Function (PDF) ดังสมการ

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-x^2/2\sigma^2} \quad (2)$$

โดยที่ σ^2 นั้น เป็นค่าความแปรปรวน (Variant) ของการสุ่ม และตัวแปร μ ในสมการนั้น เป็นค่าเฉลี่ยของข้อมูลที่ได้ออกจากการสุ่ม โดยที่ในกรณีนี้แสดงถึงสัญญาณรบกวนนั้น จะมีค่าเฉลี่ยเป็นศูนย์สาเหตุที่ใช้สัญญาณรบกวนขาว (White noise) เนื่องจากสีขาวนั้นสื่อถึงสัญญาณรบกวนแบบที่มีองค์ประกอบของความถี่เท่ากันทุกๆความถี่

3. วิธีการดำเนินการวิจัย

ในการออกแบบระบบการสื่อสารแบบไร้สายในระบบ OFDM จำเป็นต้องเข้าใจในหลักการการทำงานของระบบเสียก่อน เพื่อวิเคราะห์ปัญหาที่เกิดขึ้น จากนั้นจึงสามารถพัฒนาอุปกรณ์อื่นๆที่ใช้งานร่วมกันได้ ในการจำลองระบบนี้จะแบ่งการทำงานออกเป็น 2 ส่วนนั้น คือส่วนของการรับ-ส่งระบบ OFDM ที่มีและไม่มี การป้อนข้อมูลย้อนกลับ เพื่อปรับชดเชยสัญลักษณ์ ให้เห็นข้อแตกต่างของประสิทธิภาพในแต่ละระบบสามารถนำไปวิเคราะห์ได้ง่าย และส่วนของการประมวลผลจะใช้โปรแกรม MATLAB ในการสร้างระบบจำลองขึ้นมาโดยอ้างอิงมาจาก [6] ในภาพรวมมีรูปแบบการทำงานในระบบที่จะนำเสนอตั้งรูปที่ 2



รูปที่ 2: แบบจำลองระบบช่องการสื่อสาร

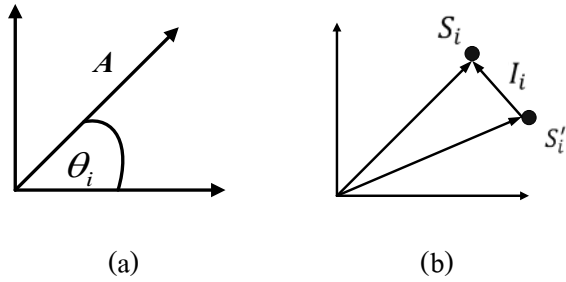
ขั้นตอนการทำงานของแบบจำลองมีดังนี้

3.1 ในตำแหน่งภาคส่งจะสร้างบิตข้อมูลที่มีรูปแบบเป็นเลขฐานสอง

3.2 ทำการมอดูเลตให้ข้อมูลอยู่ในรูปของจำนวนเชิงซ้อนที่มีความสัมพันธ์ไอและคว (IQ: Inphase Quadrature)

3.3 ปรับชดเชยสัญลักษณ์ในภาค Constellation โดยรับสัญญาณ Low information Feedback จาก Power and Phase Detector ในภาครับ ซึ่งการป้อนกลับของข้อมูลขนาดต่ำมากนี้ช่วยบอกถึงลักษณะช่องสัญญาณก่อนที่จะส่งข้อมูลข่าวสารผ่านช่องทางการสื่อสารเดิม สัญญาณที่ได้ จะผ่านลักษณะช่องสัญญาณที่เป็นพหุวิธี พร้อมทั้งสัญญาณรบกวนแบบเกาส์ขาวแบบบวก เมื่อเข้าถึงภาครับระบบ OFDM สัญญาณจะ

ผ่านไปยังอุปกรณ์แยกสัญญาณข้อมูลและสัญญาณนำร่องออกจากกัน สัญญาณนำร่องจะถูกใช้เพื่อเป็นข้อมูลป้อนกลับ



รูปที่ 3: ความสัมพันธ์ในการจัดกลุ่มสัญลักษณ์

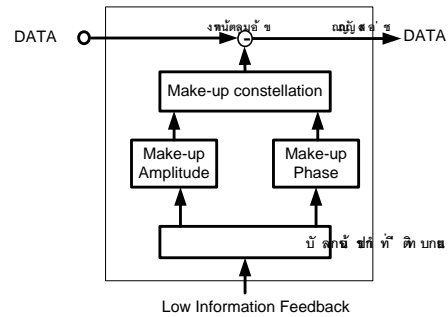
การปรับขนาดเซกกลุ่มสัญลักษณ์ จะเป็นการส่งสัญญาณนำร่องเพื่อตรวจสอบลักษณะช่องสัญญาณเป็นระยะๆ งานวิจัยนี้ได้ศึกษาโครงสร้างรวมถึงรูปแบบการรับและปรับขนาดเซกข้อมูลที่มีข้อผิดพลาด ทั้งขนาดและเฟสของข้อมูลขณะผ่านช่องสัญญาณ เมื่อภาครับทราบถึงการเปลี่ยนแปลงสภาพช่องสัญญาณแล้ว จึงทำการส่งข้อมูลไปช่องสัญญาณที่มีการปรับแก้คืบไปยังภาคส่งเพื่อจัดกลุ่มภายใต้การทำงานในระบบมอดูเลต พบว่าจากรูปที่ 3(a) สัญลักษณ์ของสัญญาณที่ถูกส่งออกมาแต่ละเวลาอยู่ในช่วงของ t จะได้สมการที่ (3)

$$S_i = Ae^{j\theta} \tag{3}$$

โดยค่าสัมประสิทธิ์ของ A เป็นค่าที่แสดงถึงขนาดของสัญญาณ และ θ แสดงถึงเฟสที่ถูกส่งออกมาในช่วงเวลา t_k วิธีการจัดกลุ่มสัญลักษณ์แบบปรับขนาดเซกนั้นจะจัดกลุ่มสัญลักษณ์ล่วงหน้าเพื่อที่ว่าเมื่อสัญลักษณ์ถึงปลายทางแล้วจะไม่มีผลกระทบของสัญญาณหรือส่งผลกับสัญลักษณ์ให้เกิดผลน้อยที่สุด ซึ่งคำนวณได้จากรูปที่ 3(b) สามารถเขียนสมการที่ (4) ได้ดังนี้

$$S'_i = S_i + \sum I_i \tag{4}$$

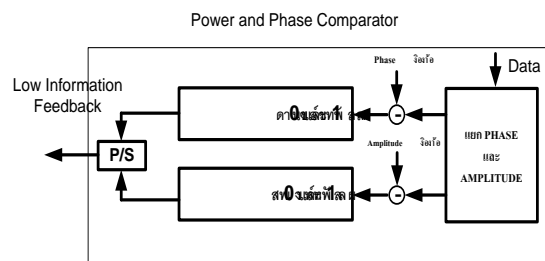
โดยสัญญาณ S'_i คือ ผลที่ได้จากการคำนวณเป็นข้อมูลที่ส่งออกจากภาคจัดกลุ่มแบบปรับขนาดเซกหลังจากได้ Low Information Feedback เป็นข้อมูลการปรับแก้กลุ่มสัญลักษณ์ส่วน I_i เป็นตัวแปรที่ถูกนำมาหักล้างจากสัญลักษณ์ที่ได้จากภาคมอดูเลตใช้ชดเชยการแทรกสอดระหว่างกลุ่มสัญลักษณ์ที่เกิดขึ้นในช่องสัญญาณที่ถูกส่งออกไปก่อนหน้านี้ดังแสดงในรูปที่ 4



รูปที่ 4: การจัดกลุ่มใหม่ที่ภาค Transmitter

3.4 ข้อมูลที่ได้ปรับแก้แล้วจะถูกส่งเข้าไปยังแบบจำลองระบบ ผู้วิจัยได้เอาแบบจำลองที่อ้างอิงมาจาก [6] และได้เพิ่มช่องสัญญาณแบบพหุวิถีอีก 5 เส้นทาง

3.5. ในภาครับเมื่อผ่านแบบจำลองระบบ จากรูปที่ 5 ข้อมูลจะถูกแยก เฟสและแอมพลิจูดเพื่อเทียบกับสัญญาณอ้างอิง ดังรูปที่ 5 จากนั้นข้อมูลที่ได้จะส่งย้อนกลับครึ่งละสองบิตเพื่อเป็นข้อมูลที่ใช้ในการปรับแก้สัญลักษณ์ต่อไป



รูปที่ 5: การเปรียบเทียบขนาดและเฟสที่ภาค Receiver

3.6 พารามิเตอร์ที่ใช้เกี่ยวข้องกับการทดลอง

เพื่อให้ผลการทดลองใกล้เคียงกับระบบการสื่อสารแบบไร้สายที่ใช้งานในปัจจุบันมากที่สุด จึงนำพารามิเตอร์ต่างๆที่เกี่ยวข้องเข้ามาประกอบในแบบจำลองระบบที่ออกแบบขึ้นมา ดังนี้

3.6.1 พารามิเตอร์ของช่องสัญญาณแบบพหุวิถี ในแบบจำลองระบบนี้เป็นส่วนประกอบที่สำคัญสำหรับการสื่อสารแบบไร้สาย ซึ่งแต่ละช่องสัญญาณจะมีลักษณะที่แตกต่างกันออก ส่งผลทำให้ข้อมูลผิดพลาด ซึ่งงานวิจัยนี้จำลองเส้นทางขึ้น 5 เส้นทาง

3.6.2 พารามิเตอร์ที่ใช้ปรับขนาดและเฟสของข้อมูลที่ต้องการส่งโดยอัตราการปรับขนาดเขยดังตารางที่ 3 โดยมี Feedback information ย้อนกลับส่งกลับมารั้งละ 2 บิต ซึ่งข้อมูลที่ย้อนกลับมาจะบอกถึงลักษณะของข้อมูลมีขนาดและเฟสที่เปลี่ยนแปลงไปจากเดิมในภาครับ

ตารางที่ 1: อัตราการปรับขนาดเขย

Feedback information	การปรับขนาดเขยสัญลักษณ์
00	ลดแอมพลิจูด 0.5 dB ลดเฟส 5°
01	ลดแอมพลิจูด 0.5 dB เพิ่มเฟส 5°
10	เพิ่มแอมพลิจูด 0.5 dB ลดเฟส 5°
11	เพิ่มแอมพลิจูด 0.5 dB เพิ่มเฟส 5°

4. ผลการดำเนินงาน

พารามิเตอร์ในระบบ OFDM ผู้วิจัยได้นำมาจาก [5-6] ซึ่งเป็นไปตามมาตรฐาน IEEE802.11a [8]

ตารางที่ 2: พารามิเตอร์ที่ใช้ในการกำหนดระบบ

No.	Parameter	Value
1.	Data	64 Kbit
2	Carrier Modulation	QAM (4PSK)
3	data sub carriers	52
4	Number of IFFT Size	64
5	Cyclic prefix length	16
6	Channel type	Rayleigh channel
7	Equalizer	Zero Forcing

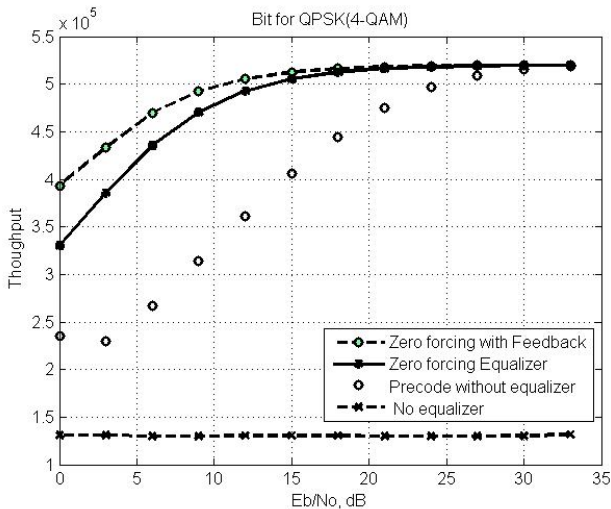
จากการทดลองตามสภาพแวดล้อมที่กำหนดขึ้น ผู้วิจัยได้ทำการทดสอบด้วยการส่งข้อมูลจำนวน 10 รอบ รอบละ 64,000 บิต โดยเป็นอิสระทางสถิติและนำผลทั้งหมดมาเฉลี่ยกัน ซึ่งใน

การทดลองในสภาพแวดล้อมนี้ ใช้พารามิเตอร์ของช่องสัญญาณแบบพหุวิถี และมีพลังงานในการส่งแต่ละสัญลักษณ์ที่เท่ากัน ค่าพลังงานจาก [4],[9] ที่ได้ยกมาใช้ในการคำนวณ เพื่อปรับแก้สัญลักษณ์เปรียบเทียบที่มีการปรับขนาดเขยขนาดและเฟสกับแบบที่มี Tomlinson-Harashima Pre-coding (THP)[10] , ไม่มี Equalizer , Pre-coding , และแบบ Equalizer เพียงอย่างเดียวของสัญญาณแบบพหุวิถี 5 วิถี

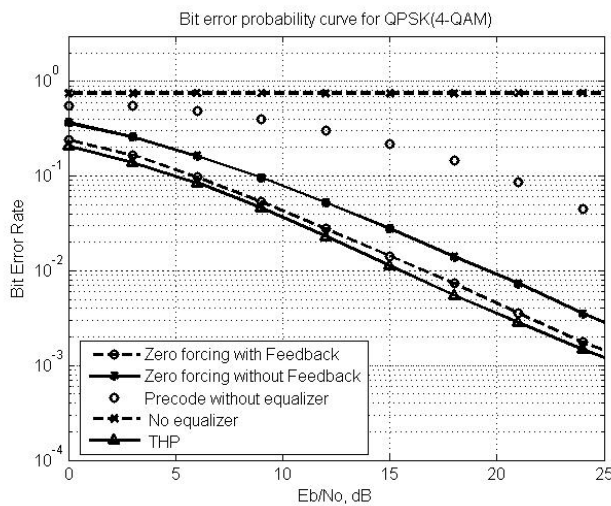
จากการทดลองจำลองการทำงานเพื่อหา Throughput และ BER ที่มีระดับพลังงานของการรับและอัตราบกววนในแต่ละสัญลักษณ์มีอัตราผิดพลาดเท่ากันดังรูปที่ 6 และการทดสอบหาอัตราผิดพลาดดังรูปที่ 7 เพื่อให้มองเห็นถึงการสะท้อนของจำนวนวิถีที่มากขึ้น เมื่อเทียบกับแบบ Equalizer เพียงอย่างเดียว ที่ $E_b/N_0 = 20$ dB ผลที่ได้คือแบบที่นำเสนอจะดีกว่าอยู่ 3 dB และเมื่อเทียบกับ THP ที่มีการคำนวณที่ซับซ้อนกว่า จะต่างกันไม่ถึง 1 dB ซึ่งระบบที่นำเสนอนี้เป็นผลทำให้ภาครับทำงานน้อยลงและลดข้อมูลผิดพลาดได้ดียิ่งขึ้น

5. สรุปผลการทดลอง

จากการทดลองระบบที่นำเสนอที่การปรับขนาดเขยกลุ่มสัญลักษณ์ใหม่ โดยเทียบกับ THP ที่เป็นการเข้ารหัสที่มีความซับซ้อนและเทียบกับภาครับที่มีเพียง Equalizer ระบบที่นำเสนอนี้มีประสิทธิภาพดีในภาพแวดล้อมที่มีการสะท้อนแบบพหุวิถี พร้อมทั้งมีความซับซ้อนที่น้อยกว่า ส่วนอัตราผิดพลาดของข้อมูลนั้นต่างกันเพียงเล็กน้อย เมื่อเทียบกับแบบ THP ซึ่งเข้ารหัสที่มีความซับซ้อนสูงกับระบบที่ได้นำเสนอที่มีความซับซ้อนน้อยกว่า โดยผู้วิจัยจะปรับปรุงรูปแบบการปรับขนาดเขยทั้งขนาดและเฟส ที่เหมาะสมพร้อมทั้งสามารถใช้ได้หลายรูปแบบ ในอนาคตต่อไป



รูปที่ 6: เปรียบเทียบประสิทธิภาพในมุมมอง Throughput ของการใช้งานแต่ละรูปแบบ



รูปที่ 7: เปรียบเทียบประสิทธิภาพ BER ภายใต้การประมาณ ช่องสัญญาณที่ไม่เป็นอุดมคติ

เอกสารอ้างอิง

[1] พรชัย ทรัพย์นิธิ, "การมัลติเพล็กซ์แบบความถี่ตั้งฉาก" ระบบสื่อสารดิจิทัล, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, หน้า 160-167, 2551

[2] ลัญฉกร วุฒิสัทติกุลกิจ, "อีควอลไรเซอร์และไดเวอร์ซิตี", หลักการระบบโทรศัพท์เคลื่อนที่. สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย, หน้า 83-90, 2542

[3] ดิสพล น้าเนียวกุล, ศักดิ์ ศรีพิมานเกียรติวัฒน์ และคณะ, "การศึกษาการใช้งานร่วมกันระหว่างการเข้ารหัส Turbo Codes กับ

Precoding", <http://www.kmitl.ac.th/dslabs/research.html> [สืบค้นเมื่อ มกราคม 2556]

[4] ชัยพร เขมะภาคะพันธ์ และ ทศพร จันทร์, "การจัดกลุ่มสัญลักษณ์แบบปรับขนาดเซย์ สำหรับการสื่อสารเคลื่อนที่บนช่องการสัญญาณที่มีการจางหายเรย์ลีแบบเลือกความถี่" *ECTI-CARD2010*, พัทยา, หน้า 192-197, 2553

[5] Sinem Coleri , Mustafa Ergen , Anuj Puri , "Channel Estimation Techniques Based on Pilot Arrangement in OFDM System" *IEEE Transactions on broadcasting*, pp. 223-229, September 2002

[6] Orlandos Grigoriadis , Srikanth Kamath, "BER Calculation Using Matlab Simulation For OFDM Transmition", *Proceeding of the International MultiConference of Engineering and Computer Scientists*, pp. 1426-1428, March 2008

[7] Ron Hranac , Bruce Currivan, "Digital Transmission Carrier-to-Noise Ratio, Signal-to-Noise Ratio and Modulation Error Ratio"[online] 2006. available from: http://www.cisco.com/c/en/us/products/collateral/video/ubr10000-series-universal-broadband-routers/prod_white_paper0900aecd805738f5 [2013, February 6].

[8] Anibal Luis Intini "Orthogonal Frequency Division Multiplexing for Wireless Networks", *Electrical and Computer Engineering Department*, pp.18-20, 2010.

[9] Robert Fischer, "Tomlinson-Harashima Precoding in Space-Time Transmission for Low-Rate Backward Channel", *International Zurich Seminar* , pp. 1 – 6 , February 2002

[10] Robert Fischer, "Precoding Scheme", *Precoding and Signal shaping for Digital Transmittion, John Wiley & Son Inc* , pp. 127-135, 2004

การประเมินประสิทธิภาพและคุณภาพของการให้บริการสื่อประสมบนเครือข่ายไร้สาย

Performance and Quality Evaluation of Multimedia Applications over Wireless LAN

ปานวิทย์ ฐะนุติ (Panwit Tuwanut)¹ และ ภาคภูมิ พรประทานเวช (Parkpoom Pornpratanwech)²
 คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง เขตลาดกระบัง กทม.
¹panwit@it.kmitl.ac.th, ²xkungkub001@gmail.com

บทคัดย่อ

ในงานวิจัยนี้ เป็นการประเมินหาประสิทธิภาพ ของการ ให้บริการสื่อประสม ในเครือข่ายไร้สาย Wi-Fi สำหรับกลุ่ม ธุรกิจขนาดเล็กถึงขนาดกลาง ด้วยโปรแกรม OPNET โดย มุ่งเน้นไปที่การให้บริการ VoIP และ Video Conference เป็น หลัก โดยมีการประเมินประสิทธิภาพเครือข่ายแบบพื้นฐานที่ใช้ งานแบบปกติในสถานะต่างๆ เพื่อหาค่า End to End delay, Packet receive และ Packet drop ต่อมาจึงทำการจำลองการ ปรับปรุงคุณภาพการ ให้บริการ สองแนวทาง ได้แก่ การจัดสรร ทรัพยากร และการจัดลำดับความสำคัญของคิว เพื่อนำผลลัพธ์ที่ ได้ไปปรับปรุงและวางแผนการออกแบบระบบเครือข่าย ให้มี ประสิทธิภาพการ ให้บริการที่ดียิ่งขึ้น

คำสำคัญ: ประสิทธิภาพเครือข่าย การส่งเสียงผ่านเครือข่าย การจำลองโดยโปรแกรม OPNET

Abstract

This paper presents a performance evaluation of multimedia application over wireless LAN in a small-medium business by using OPNET simulation. The main object of this simulation is focused on VoIP and Video conference application. The parameters, End to End delay, Packet receive and Packet drop are used to analyze a performance of a network. Next, the QoS is applied to improve a performance in two methods, resource reservation and priority queues. The results of this simulation are employed to design a network which has a better performance.

Keywords: Network performance, VoIP, OPNET, Video conference

1. บทนำ

จากการเติบโตของเครือข่ายอินเทอร์เน็ตที่มีความเร็วในการ ให้บริการสูงขึ้น ส่งผลทำให้กลุ่มธุรกิจขนาดเล็ถึงขนาดกลาง มีการติดตั้งระบบเครือข่ายใช้งานในองค์กร โดยรองรับการใ้ งานเบื้องต้น อาทิเช่น การใช้งาน Http, Mail เป็นต้น และต่อมา แนวคิดในการรวมระบบต่างๆ มารวมกันบนเครือข่าย อินเทอร์เน็ต ที่เรียกว่าเป็น Convergence Network ทำให้ระบบ เครือข่ายในองค์กรมีการ ให้บริการที่หลากหลาย โดยแบ่งเป็น ประเภทได้สองแบบหลักๆ คือ แบบ Real Time อาทิเช่น VoIP, Video Conference และแบบ Non Real Time อาทิเช่น Http, Ftp, Mail เป็นต้น [2] ซึ่งการที่มีบริการที่หลากหลายจึงส่งผล กระทบต่อคุณภาพการ ให้บริการ (QoS)

จากการสืบค้นงานวิจัย ได้พบงานวิจัย Yan Chen และคณะ [2] และงานของ Ajith Kumar V. และคณะ [3] ได้ศึกษาการ กำหนดค่าพารามิเตอร์ต่างๆ ที่เหมาะสมของแต่ละ Application ต่อมาในงานของ Dr. Hussein A [4] ได้กล่าวถึงเทคนิคการทำ QoS แบบต่างๆ ของการ ให้บริการ VoIP บนเครือข่าย LAN และต่อมาก็ได้มีงานวิจัยเกี่ยวกับการหาประสิทธิภาพ VoIP บน เครือข่าย Wireless LAN อาทิเช่นงานของ A. Al-Naamany [5] ต่อมาในงานของ Fabien Houéto [6] ก็เป็นการหา ประสิทธิภาพของการ ให้บริการ Voice และ Video บนระบบ เครือข่าย LAN ซึ่งจากการสืบค้นก็พบข้อจำกัดว่า ยังไม่มีงานใด ที่ทำการหาประสิทธิภาพของทั้งบริการ VoIP และ Video Conference พร้อมกับบริการอื่นๆ บนเครือข่ายไร้สาย ดังนั้น ในงานวิจัยนี้ จึงเป็นการประเมินค่าประสิทธิภาพการ ให้บริการ VoIP และ Video Conference บนเครือข่ายไร้สาย พร้อมกับ บริการอื่นๆ เช่น Http, Ftp, Mail สำหรับกลุ่มธุรกิจขนาดเล็กถึง ขนาดกลาง ที่มีการติดตั้งระบบเครือข่ายไร้สายในองค์กร โดย จำลองสถานการณ์ระบบเครือข่ายของกลุ่มธุรกิจที่ใช้งานจริง

เพื่อหารูปแบบการทำ QoS ที่เหมาะสม ซึ่งในการจำลองนี้ ได้ทำการจำลองโดยใช้โปรแกรม OPNET

โครงสร้างของงานวิจัยนี้ได้แบ่งเป็น 5 ส่วน ซึ่งส่วนแรกคือ บทนำ ส่วนถัดมาคือ ทฤษฎีที่เกี่ยวข้อง ในส่วนที่สามกล่าวถึง การจำลองเครือข่าย และการตั้งค่า ในส่วนที่สี่กล่าวถึงผลการจำลอง โดยโปรแกรม OPNET และการอภิปรายผลการจำลอง และสรุปผลกล่าวสรุปในส่วนสุดท้าย

2. ทฤษฎีที่เกี่ยวข้อง

2.1 Quality of Service (QoS)

ในเครือข่ายการสื่อสาร ที่รองรับการให้บริการประเภทต่างๆของผู้ใช้บริการจะทำงานได้ดี เมื่อเครือข่ายนั้นสามารถตอบสนองความต้องการของผู้ใช้บริการ โดยรับประกันคุณภาพการบริการโดยมีความน่าเชื่อถือ ซึ่งมีค่าเวลาหน่วงและคุณภาพการบริการได้ตามที่กำหนด โดยมีอยู่หลายวิธีอาทิเช่น

- First-In-First-Out (FIFO) เป็นระบบแถวคอย ที่ไม่มีการระบุประเภทของสิทธิในการเข้าใช้ โดยใช้หลักมาถึงก่อนได้รับบริการก่อน ซึ่งระบบแถวคอยประเภทนี้มีประสิทธิภาพ และเหมาะสมสำหรับการเชื่อมโยงที่มีความจุสูง ไม่ต้องการความล่าช้าหรือความคับคั่งมาก

- Priority Queuing (PQ) เป็นระบบแถวคอย ที่มีการกำหนดลำดับความสำคัญของประเภทข้อมูลที่เข้ามาในแถวคอย ให้มีสิทธิในการสื่อสารตามลำดับสิทธิที่กำหนด โดยในที่นี้แบ่งเป็นสี่ระดับ คือ High, Medium, Normal และ Low

- Weighted Fair Queuing (WFQ) เป็นระบบแถวคอย ที่ใช้การจัดสรรแบนวิดท์ ให้กับข้อมูลแต่ละประเภท เพื่อป้องกันข้อมูลที่มีปริมาณสูงครอบครองแบนวิดท์ในการสื่อสาร

2.2 พารามิเตอร์ที่สำคัญในการควบคุม QoS [2]

1. ค่า Jitter หรือ Packet Delay Variation คือ ค่าความแปรปรวนของเวลาที่ packet ใช้ในการเดินทางมาถึงปลายทาง

2. ค่า Packet End to End delay คือ ค่าผลต่างของเวลาตั้งแต่ packet ถูกส่งจากต้นทาง และปลายทางได้รับ packet

3. ค่า Packet Loss คือ ค่าการสูญหายของ packet ที่เดินทางไม่ถึงปลายทาง วัดได้จากจำนวน packet ที่ส่งจากต้นทางลบกับจำนวน packet ที่ปลายทางได้รับ

4. ค่า MOS คือ การวัดคุณภาพเสียง โดย ITU ได้มีวิธีการทดสอบโดยกำหนดข้อความที่ใช้ทดสอบ เรียกว่า Harvard Sentences ซึ่งมีการออกเสียง ในช่วงความถี่เสียงทั้งหมดที่พบในบทสนทนาทั่วไป โดยบันทึกด้วยความละเอียดขนาด 16 บิต ตัวอย่างเช่น You will have to be very quiet, There was nothing to be seen เป็นต้น

5. ค่า Response Time Expected by user คือ ค่าคาดหวังต่อการตอบสนองการให้บริการ ของผู้รับบริการ เช่น ค่าคาดหวังในการตอบสนองการบริการ Http อยู่ในช่วงเวลา 2-5 วินาที [2]

2.3 การส่งเสียงผ่านเครือข่ายไอพี (VoIP)

ในปัจจุบันเทคโนโลยีด้านอินเทอร์เน็ต มีพัฒนาไปไกลและมีความเร็วสูง จึงมีแนวคิดในการรวมระบบต่างๆ มารวมกันบนเครือข่ายอินเทอร์เน็ต ที่เรียกว่าเป็น Convergence Network ซึ่งระบบโครงข่ายโทรศัพท์พื้นฐาน (PSTN) ก็ถูกพัฒนามาอยู่บนเครือข่ายอินเทอร์เน็ต เรียกว่า Voice over IP (VoIP) ซึ่งในการสื่อสาร VoIP นิยมใช้โปรโตคอล SIP และ สถาบัน ITU กำหนดค่าพารามิเตอร์มาตรฐานของ VoIP ในการควบคุมคุณภาพการสื่อสาร ดังตารางที่ 1

ตารางที่ 1: ค่ามาตรฐาน QoS ของ VoIP

	Average Quality	Ideal Quality
Jitter	< 60 ms	< 20 ms
End-to-end Delay	< 150 ms	< 50 ms
Packet Loss Rate	< 5%	< 1%

Quality Scale	Mean Opinion Score (MOS)
Excellent	5
Good	4
Fair	3
Poor	2
Bad	1

2.4 การประชุมทางไกล (Video Conference)

ระบบประชุมทางไกล เป็นระบบที่ผสมผสานระหว่าง ข้อมูลภาพ และเสียง ให้เปรียบเสมือนมีการประชุมอยู่ในห้องเดียวกัน และในการใช้งาน Video Conferencing บนเครือข่าย จะมีการกำหนดค่าพารามิเตอร์มาตรฐานในการควบคุมคุณภาพของการสื่อสารไว้ ดังตารางที่ 2

ตารางที่ 2: ค่ามาตรฐาน QoS ของ Video Conference

Traffic Class	Technology Attributes	QoS Metrics							
		Timeliness			Precisioness			Accuracy	
		Response time Expected by Users	Delay (ms)	Jitter (ms)	Data Rate (bps)	Required Bandwidth (bps)	Loss Rate	Error Rate	
Video Conferencing	Real Time and Symmetric	Lip-synch: <100 ms	<150	<400				<0.01%	< 0.01%
Coding Standard									
		H.320			64-1920K	80K-2M			
		H.323			64X K	80X K			
		H.324			<64K	<80K			
Network Capacity									
Link: Refer to Appendix I					Router: Refer to Appendix II				

2.5 File Transfer Protocol (FTP)

เป็นโปรโตคอลที่ใช้สำหรับแลกเปลี่ยนและจัดการไฟล์บนเครือข่ายที่ซีพี/ไอพี ในการใช้งาน Ftp จะมีค่าพารามิเตอร์ที่ใช้วัดประสิทธิภาพในการทำงานผ่านเครือข่ายคือค่า Response Time Expected by user โดยค่าคาดหวังที่ตอบสนองต่อผู้ใช้บริการอยู่ในช่วง 2-5 วินาที [2]

2.6 Hypertext Transfer Protocol (HTTP)

เป็นโปรโตคอล ที่ในการร้องขอและการตอบรับระหว่างเครื่องลูกข่ายกับเครื่องแม่ข่ายที่เก็บข้อมูลของเว็บไซต์ไว้ ค่าพารามิเตอร์ที่ใช้วัดประสิทธิภาพในการทำงานผ่านเครือข่ายคือค่า Response Time Expected by user เพื่อวัดเมื่อผู้ใช้งานต้องการโหลดหน้าเว็บไซต์ต่างๆขึ้นมาชม โดยค่าคาดหวังที่ตอบสนองต่อผู้ใช้บริการควรอยู่ในช่วง 2-5 วินาที [2]

2.7 ไปรษณีย์อิเล็กทรอนิกส์ (Electronic Mail: E-mail)

เป็นวิธีการแลกเปลี่ยนข้อความแบบดิจิทัล โดยข้อความนั้นประกอบด้วยเนื้อหา ที่อยู่ของผู้ส่ง และที่อยู่ของผู้รับ โดยผ่านเครื่องแม่ข่ายให้บริการ E-mail (Mail Server) โดยค่าพารามิเตอร์ที่ใช้วัดประสิทธิภาพในการทำงานผ่านเครือข่ายคือค่า Response Time Expected by user โดยค่าคาดหวังที่ตอบสนองต่อผู้ใช้บริการอยู่ในช่วง 2-5 วินาที [2]

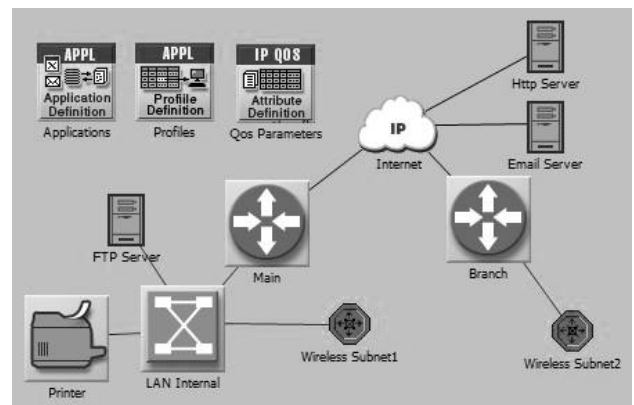
3. การจำลองเครือข่าย และการตั้งค่า

ในการวัดประสิทธิภาพการให้บริการ VoIP จะใช้พารามิเตอร์หลักๆ ดังนี้ Voice Jitter, End to End delay, Packet Loss Rate และ MOS Value สำหรับการวัดประสิทธิภาพของการให้บริการ Video Conference จะใช้พารามิเตอร์หลักๆ ได้แก่ End to End delay, Packet Loss Rate, Packet Delay และค่า Throughput และในการให้บริการ Http และ Ftp จะใช้ค่า Response Time Expected by user

3.1 เหตุการณ์ที่ใช้ในการจำลอง

มีการกำหนดสถานการณ์จำลองให้คล้ายคลึงกับการใช้งานจริงในกลุ่มธุรกิจขนาดเล็กลงถึงขนาดกลางที่มีสำนักงานสองที่ โดยที่สาขาหลัก มีการให้บริการ Ftp Server และ Printer Server และมีเครื่องลูกข่ายในสำนักงานเชื่อมต่อ Wi-Fi จำนวน 40 เครื่อง ผ่าน Access Point 1 ตัว ส่วนที่สาขาย่อยมีเครื่องลูกข่ายเชื่อมต่อ Wi-Fi จำนวน 30 เครื่อง ผ่าน Access Point 1 ตัว และทั้งสองสาขา มีการใช้งาน Web Server และ Email Server ที่ตั้งอยู่นอกสำนักงาน และการทำ QoS จะทำที่บนสายเชื่อมต่อระหว่าง Internet กับเราเตอร์ทั้งสองตัว (Main และ Branch) โดยอุปกรณ์ที่มีการเชื่อมต่อเข้ากับ Internet จะใช้สาย PPP DS1 (1.544 Mbps) ในการเชื่อมต่อ ส่วนอุปกรณ์อื่นๆที่เชื่อมต่อกันเป็นเครือข่ายภายในจะใช้สาย Ethernet 100BaseT ในการเชื่อมต่อ

- แสดง Topology ที่ใช้ในการจำลองดังภาพที่ 1



ภาพที่ 1: Topology ที่ใช้ในการจำลอง

- โดยกำหนดค่า Profile ของ Application [1] ดังตารางที่ 3

ตารางที่ 3: การกำหนดค่า Profile ของ Application

Application	Load	Inter Arrival Time (sec)
Http	High Browsing	60
Ftp	High Load	360
Email	High Load	360
Printer	Color Print	360
VoIP	PCM Quality Speech	900
Video Conference	High Resolution	900

ในการจำลอง มีการกำหนดจำนวนเครื่องคอมพิวเตอร์ที่ใช้
งานบริการต่างๆ ดังตารางที่ 4

ตารางที่ 4: การกำหนดจำนวนเครื่องคอมพิวเตอร์ที่ใช้บริการต่างๆ

Applications	สำนักงาน (Main)	สาขาย่อย (Branch)
Http	15	10
Ftp	5	5
Email	5	5
Printer	5	-
VoIP	5	5
Video Conference	5	5

ในการทดลองที่มีการทำ QoS ได้ใช้วิธีการทำอยู่ 3 แบบ ซึ่ง
มีการตั้งค่าในโปรแกรม ดังตารางที่ 5

ตารางที่ 5: การกำหนดกลไกที่ใช้ในการทำ QoS

QoS Scheme	QoS Profile
FIFO	FIFO Profile (Default)
Priority Queuing	Type of Service Based
Weight Fair Queuing	Bandwidth Management

โดย Priority Queuing มีการกำหนดอยู่ 4 ระดับ ซึ่งแต่ละ
ระดับจะแบ่งตาม Type of Service ของแต่ละแอปพลิเคชันใน
การบริการ ดังตารางที่ 6

ตารางที่ 6: การกำหนด ToS ของ Priority Queuing แต่ละระดับ

Priority Level	Type of Service (ToS)
0 (Low)	Best Effort (0), Background (1)
1 (Normal)	Standard (2), Excellent Effort (3)
2 (Medium)	Streaming Multimedia (4), Interactive Multimedia (5)
3 (High)	Interactive Voice (6), Reserved (7)

และในการทำ Weight Fair Queuing โดยใช้ Bandwidth
Management มีการกำหนดค่าโดยดูจาก Type of Service ของ
แอปพลิเคชัน ดังตารางที่ 7

ตารางที่ 7: การกำหนดน้ำหนักของ ToS ที่ใช้ใน Weight Fair Queuing

Bandwidth Management	Type of Service
30	Best Effort (0)
40	Interactive Multimedia (5)
30	Interactive Voice (6)

4. ผลการจำลองโดยโปรแกรม OPNET และการ อภิปรายผลการจำลอง

จากการจำลองสถานการณ์ดังกล่าวที่ 1 มาทดลองได้ผลว่า
เมื่อไม่มีการทำ QoS ไม่ว่าจะใช้บริการ VoIP คุณภาพระดับใดก็
ตาม (Low, GSM, PCM) จะได้ค่าของพารามิเตอร์ที่ใช้วัด
ประสิทธิภาพออกมาไม่อยู่ในเกณฑ์มาตรฐานที่มีกำหนดไว้ ทำ
ให้ได้การบริการที่ไม่มีประสิทธิภาพที่ดี ดังตารางที่ 8

ตารางที่ 8: แสดงค่าพารามิเตอร์ของ VoIP ต่างคุณภาพ

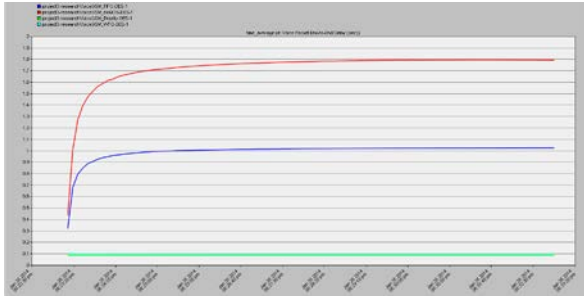
Quality	Jitter (sec)	End to End Delay (sec)	MOS Value	Data Loss Rate	FTP Download Response Time (sec)	HTTP Page Response Time (sec)
PCM	0.001	1.25	3.7	49.44%	110	32
GSM	0.0001	1.8	3.58	42.22%	20	30
Low (G723.1)	0.0001	2.1	2.59	33.33%	9	27

ดังนั้นจึงมีการจำลองสถานการณ์เปรียบเทียบวัด
ประสิทธิภาพของบริการ Voice ที่คุณภาพระดับกลาง (GSM)
ทั้งแบบไม่มีการทำ QoS กับแบบทำ QoS ด้วยกลไกต่างๆขึ้นมา
เพื่อทำการเปรียบเทียบ ซึ่งได้ผลว่าการทำ QoS ด้วยกลไก
Priority Queuing มีประสิทธิภาพที่ดีที่สุด ซึ่งได้ผลออกมา
สอดคล้องตรงตามกับทฤษฎี เนื่องจากมีเกณฑ์ของ
ค่าพารามิเตอร์ในการวัดออกมามีค่าที่ดีที่สุด โดยค่า End to End
Delay เท่ากับ 0.1 วินาที, MOS Value เท่ากับ 3.58, Data Loss
Rate เท่ากับ 0.1% ซึ่งค่าพารามิเตอร์ดังกล่าวได้สอดคล้องตาม
เกณฑ์มาตรฐานที่ได้กำหนดไว้ ดังตารางที่ 9

ตารางที่ 9: แสดงค่าพารามิเตอร์ของ VoIP ที่ทำ QoS

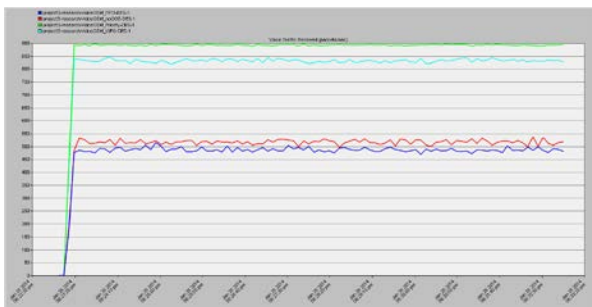
	Jitter (sec)	End to End Delay (sec)	MOS Value	Data Loss Rate	FTP Download Response Time (sec)	HTTP Page Response Time (sec)	
ไม่ทำ QoS	0.001	1.8	3.58	41.67%	20	32	
ทำ QoS	FIFO	0.001	1.12	3.58	46.11%	12	27
	PQ	0	0.1	3.58	0.1%	1.2	0.1
	WFQ	0	0.1	3.58	6.67%	10	5.9

ต่อมาในส่วนของผลการทดลองที่แสดงเป็นกราฟกำหนดให้แกนแนวนอนเป็นแกนของค่าเวลาที่ใช้ในการทดลองนั้น และแกนแนวตั้งเป็นแกนของค่าพารามิเตอร์ที่ใช้วัดประสิทธิภาพหรือคุณภาพในการทดลองนั้น



ภาพที่ 2: กราฟแสดงพารามิเตอร์ End to End Delay

ดังภาพที่ 2 เห็นได้ว่าพารามิเตอร์ Packet End to End Delay ของ VoIP ที่ทำ QoS กลไก PQ และ WFQ มีค่าน้อยที่สุดคือ 0.1 วินาที รองลงมาคือ ทำ QoS ด้วยกลไก FIFO มีค่า 1.12 วินาที และสุดท้ายคือ ที่ไม่ได้ทำ QoS ซึ่งมีค่า 1.8 วินาที



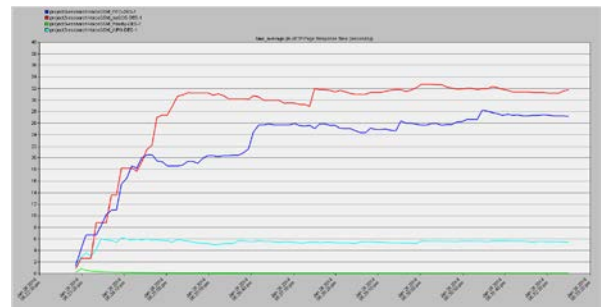
ภาพที่ 3: กราฟแสดงพารามิเตอร์ Packet Received

ดังภาพที่ 3 เป็นการแสดงค่าของพารามิเตอร์ Packet Received ของ VoIP ที่ทำ QoS และไม่ได้ทำ QoS โดยที่ทุกแบบจำลองจะมี Packet Sent จำนวนเท่ากันคือ 900 packet/sec ดังกราฟจะเห็นว่าผลที่ทำ QoS กลไก PQ ทำให้ปลายทางได้รับ Packet มากที่สุดโดยมี Packet Loss ประมาณ 0.1% รองลงมาคือ QoS กลไก WFQ โดยได้รับ Packet ประมาณ 840 packet/sec จึงมี Packet Loss อยู่ที่ประมาณ 6.67% รองลงมาคือ แบบไม่ทำ QoS โดยมี Packet Loss อยู่ที่ประมาณ 41.67% และสุดท้ายคือ ทำ QoS แบบ FIFO โดยมี Packet Loss อยู่ที่ประมาณ 46.11%



ภาพที่ 4: กราฟแสดงพารามิเตอร์วัดประสิทธิภาพของ FTP

ดังภาพที่ 4 เป็นการแสดงค่าของพารามิเตอร์ Ftp Download Response Time ที่มีการจำลองสภาพแวดล้อมอยู่บนเครือข่ายที่มีบริการ VoIP ที่ทำ QoS และไม่ได้ทำ QoS เพื่อบ่งบอกประสิทธิภาพของบริการ Ftp ซึ่งจากผลจะได้ว่า Ftp Download Response Time จะมีค่าดีสุดบนสภาพแวดล้อมที่ทำ QoS แบบ PQ รองลงมาคือ QoS แบบ WFQ ต่อมาคือ QoS แบบ FIFO และสุดท้ายคือ แบบไม่ทำ QoS โดยมีค่าพารามิเตอร์เป็น 1.2 วินาที, 10 วินาที, 12 วินาที และ 20 วินาทีตามลำดับ



ภาพที่ 5: กราฟแสดงพารามิเตอร์วัดประสิทธิภาพของ HTTP

ดังภาพที่ 5 เป็นการแสดงค่าของพารามิเตอร์ Http Page Response Time ที่มีการจำลองสภาพแวดล้อมแบบเดียวกันกับภาพที่ 4 ที่เปรียบเทียบทั้ง 4 สถานการณ์ เพื่อบ่งบอกประสิทธิภาพของบริการ Http จากผลทดลองจะได้ว่า Http Page Response Time จะมีค่าดีสุดบนสภาพแวดล้อมที่ทำ QoS แบบ PQ รองลงมาคือ QoS แบบ WFQ ต่อมาคือ QoS แบบ FIFO และสุดท้ายคือ แบบไม่ทำ QoS โดยมีค่าพารามิเตอร์เป็น 0.1 วินาที, 5.9 วินาที, 27 วินาที และ 32 วินาทีตามลำดับ

ส่วนของการจำลองสถานการณ์เปรียบเทียบประสิทธิภาพของบริการ Video Conference ทั้งแบบไม่มีการทำ QoS กับแบบทำ QoS ได้ผลออกมาคือ ค่าของพารามิเตอร์ที่ใช้วัด

ประสิทธิภาพแสดงให้เห็นว่าการทำ QoS ไม่ได้ทำให้ได้ประสิทธิภาพในการบริการที่ดีกว่าแบบไม่ได้ทำ QoS เนื่องจากการส่ง Packet ข้อมูลของบริการ Video Conference จะมีขนาดใหญ่ ทำให้มีโอกาสเกิดการสูญหายของข้อมูลก่อนไปถึงปลายทางได้สูง ยิ่งสถานะที่เป็นช่วงที่มีการใช้บริการหลายๆ อย่างบนเครือข่ายในเวลาเดียวกันก็ยิ่งส่งผลกระทบต่อประสิทธิภาพในการใช้งาน ดังตารางที่ 10

ตารางที่ 10: แสดงค่าพารามิเตอร์ของ Video Conference ที่ทำ QoS

		Packet Delay Variation (sec)	End to End Delay (sec)	Data Loss Rate
ไม่ทำ QoS		0	0.018	25%
ทำ QoS	FIFO	0	0.018	36%
	PQ	0	0.018	36%
	WFQ	0.1	0.03	73.5%

5. สรุปผล

จากการทดลองเพื่อหาประสิทธิภาพของทั้งบริการ VoIP และ Video Conference พร้อมกับบริการอื่นๆ บนเครือข่ายไร้สาย สามารถสรุปผลได้ดังนี้

- VoIP ที่เป็นแบบ PCM จะมีค่าพารามิเตอร์ Jitter และ Packet End to End Delay ดีกว่าแบบ GSM และ G723.1 แต่ถ้าดูจากค่าพารามิเตอร์ของ Packet Loss Rate จะเห็นว่า PCM นั้นมีอัตราที่ต่ำกว่าแบบอื่นและมีค่าเกินกว่าเกณฑ์มาตรฐานที่กำหนดไว้ เนื่องจากว่าที่ทำการจำลองสถานการณ์นี้แสดงให้เห็นว่าเมื่อไม่มีการทำ QoS ให้กับบริการ จึงส่งผลให้ประสิทธิภาพบริการออกมาได้ไม่ดี

- การที่บริการ VoIP ที่จำลองในสถานการณ์ที่ไม่มีการทำ QoS แสดงให้เห็นว่าประสิทธิภาพที่ไม่ดี จึงได้เปรียบเทียบกับกรณีที่ทำ QoS ในกลไกต่างๆ และได้ผลลัพธ์ออกมาว่าการบริการ VoIP ที่ทำ QoS ด้วยกลไก PQ จะทำให้ได้บริการที่มีประสิทธิภาพดีที่สุด และได้ค่าพารามิเตอร์ที่วัดคุณภาพอยู่ในเกณฑ์มาตรฐานที่ได้กำหนดไว้

- Video Conference ที่เปรียบเทียบประสิทธิภาพโดยการจำลองสถานการณ์ว่าแบบไม่ทำ QoS กับทำ QoS กลไกต่างๆ เห็นได้ว่าสถานการณ์แบบไม่ได้ทำ QoS ได้ค่าพารามิเตอร์ออกมาพอกันหรือดีกว่าสถานการณ์ที่มีการทำ QoS

- บริการอื่นๆ อาทิเช่น Http, Ftp, Mail ที่เป็นบริการแบบ Non Real Time ซึ่งมีการกำหนด ToS ที่ต่ำกว่าบริการแบบ Real Time ทำให้ได้ประสิทธิภาพออกมาไม่ดีในกรณีที่มีการใช้งานบริการพร้อมกันในเวลาเดียวกันกับบริการแบบ Real Time และมีการทำ QoS ด้วย เพราะการทำ QoS จะเน้นจัดการคุณภาพที่บริการ Real Time เป็นสำคัญ

แนวทางการวิจัยต่อไปในอนาคต คือ ทำการจำลองสถานการณ์อื่นๆ ที่มีการใช้งานอยู่จริงแบบต่างๆ เพิ่มเติมเพื่อวิเคราะห์ว่าผลที่ได้จากการจำลองในการวิจัยนี้ เมื่อเปรียบเทียบกับผลการทดลองกับสถานการณ์แบบอื่นแล้วจะได้ผลออกมาเหมือนกันหรือต่างกันอย่างไร

เอกสารอ้างอิง

- [1] E. Aboelela and Morgan Kaufmann, "Network Simulation using OPNET," Material for Laboratory, City University London, MSc in Telecommunication and Networks, Nov, 2006.
- [2] Yan Chen, Toni Farley and Nong Ye, "QoS Requirements of Network Applications on the Internet," IOS Press and the authors, pp. 55-76, 2004.
- [3] Ajith Kumar V. and Sheela Ganesh Thorenoor, "Analysis of IP Network for different Quality of Service," International Symposium on Computing, Communication and Control (ISCCC 2009), pp. 79-84, 2009.
- [4] Dr. Hussein A. Mohammed, Dr. Adnan Hussein Ali and Hawraa Jassim Mohammed, "The Affects of Different Queuing Algorithms within the Router on QoS VoIP application Using OPNET," International Journal of Computer Networks & Communications (IJCNC), vol. 5, no. 1, pp. 117-124, January, 2013.
- [5] A. Al-Naamany, H. Bourdoucen and W. Al-Menthari, "Modeling and Simulation of Quality of Service in VoIP Wireless LAN," Journal of Computing and Information Technology - CIT 16, vol. 2, pp. 131-142, 2008.
- [6] Fabien Houéto and Samuel Pierre, "Quality of service and performance issues in multiservice networks subject to voice and video traffics," Science Direct, Computer Communications 28th, pp. 393-404, 2005.

เครือข่ายเซนเซอร์ไร้สาย IEEE802.15.4 แบบหลายเกตเวย์

Multiple Gateways System for Wireless Sensor Network IEEE802.15.4

ชนะเดช แยมคลี (Chanadej Yamclee)¹ และชัยพร เขมะภตะพันธ์ (Chaiyaporn Khemapatapan)²

สาขาวิชาวิศวกรรมคอมพิวเตอร์และโทรคมนาคม คณะวิศวกรรมศาสตร์

มหาวิทยาลัยธุรกิจบัณฑิตย์

¹y_chanadej@hotmail.com, ²chaiyaporn@dpu.ac.th

บทคัดย่อ

การทำงานของเครือข่ายเซนเซอร์ไร้สายโดยทั่วไปจะมีการส่งข้อมูลที่ได้จากโหนดเซนเซอร์ไปยังโหนดโคออดิเนเตอร์ที่ทำหน้าที่เป็นเกตเวย์เชื่อมต่อกับเครือข่ายภายนอกที่มีอยู่เพียง 1 โหนด เพื่อส่งต่อข้อมูลออกจากเครือข่ายไปยังปลายทางที่ต้องการ อย่างไรก็ตามเมื่อโหนดที่เป็นเกตเวย์เกิดการเสียหายจะทำให้ไม่สามารถส่งข้อมูลออกนอกเครือข่ายนั้นได้ งานวิจัยนี้ได้นำเสนอวิธีการเพิ่มโหนดที่เป็นเกตเวย์จำนวน 2 โหนด และประยุกต์โปรโตคอลการค้นหาเส้นทางแบบ AODV กับเครือข่ายไร้สายตามมาตรฐาน IEEE802.15.4 ร่วมกับการเปลี่ยนเกตเวย์ที่ใช้เป็นโหนดปลายทางในกรณีที่โหนดที่เป็นเกตเวย์โหนดที่ตั้งต้นไม่สามารถทำงานได้

ผลการทดสอบโดยการจำลองการทำงานผ่านโปรแกรม Network Simulator 2 (NS-2) พบว่าค่า Throughput ของเครือข่ายเซนเซอร์ไร้สายที่เสนอมีค่าสูงขึ้นเฉลี่ย 1.35 Kbps และค่า Packet Delivery Ratio (PDR) มีค่าสูงขึ้นเฉลี่ย 5.39% เมื่อเทียบกับเครือข่ายเซนเซอร์ไร้สายแบบดั้งเดิม

คำสำคัญ: เครือข่ายเซนเซอร์ไร้สาย AODV หลายเกตเวย์

Abstract

The operation of wireless sensor network (WSN) is normally forwarding data from sensor node to gateway node or coordination node which connects to another network. The gateway node will subsequently transfer the data to the specific destination. When the gateway node has been damaged, the data cannot be sent out. Thus, this research presents a method by adding 2 gateway nodes and applying the routing protocol AODV over

IEEE802.15.4 WSN including gateway changing when a default gateway node comes to shutdown.

The results from the simulation using NS2 program found that throughput and Packet Delivery Ratio (PDR) of the proposed WSN increase averagely 1.35 Kbps and 5.39%, respectively, in comparison with conventional WSN.

Keyword: AODV, Wireless Sensor Network, Multiple Gateway, IEEE802.15.4

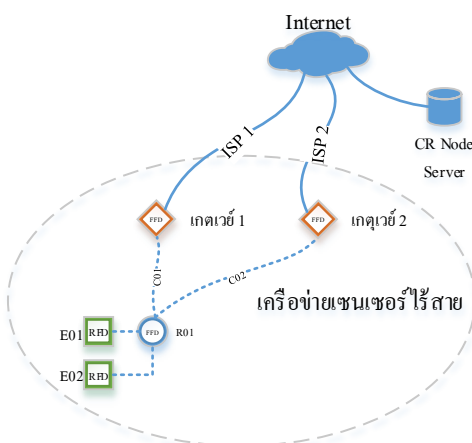
1. บทนำ

ในปัจจุบันเครือข่ายเซนเซอร์ไร้สายหรือ WSN ถูกนำมาใช้งานจริงในภาคสนามอย่างแพร่หลายทั่วโลก ในประเทศไทย มีการนำมาใช้ในงานด้านการสำรวจทางธรณีวิทยาเช่น การตรวจจับการเดินดินถล่ม งานด้านการทหาร งานด้านการเกษตร เช่นการตรวจวัด อุณหภูมิ ความชื้นในดิน [1], [2] ในการนำเอาตัวเครือข่ายเซนเซอร์ไร้สายมาประยุกต์ใช้ในงานด้านต่างๆ ก็เพื่อการเก็บข้อมูลที่ต้องการที่หน้างาน และส่งข้อมูลที่ได้เข้ามาที่ส่วนกลางเพื่อประมวลผล ซึ่งโดยโครงสร้างของ WSN จะมีการส่งข้อมูลจากโหนดเซนเซอร์มายังโหนดที่เป็นเกตเวย์ผ่านโหนดเรดเดอร์ ซึ่งโหนดปลายทางที่เป็นโหนดโคออดิเนเตอร์ซึ่งทำหน้าที่เป็นเกตเวย์จะมีอยู่เพียงแ่โหนดเดียวภายในเครือข่าย [3] ดังนั้นเมื่อโหนดเกตเวย์ได้รับความเสียหายก็จะไม่สามารถรับและส่งข้อมูลที่ต้องการ การออกนอกเครือข่ายได้ จนทำให้ข้อมูลเกิดการสูญหายสาเหตุเนื่องจากปัญหาเรื่องพลังงานและขนาดของหน่วยความจำที่มีขนาดจำกัด

ถึงแม้ว่าจะมีการแก้ปัญหาโดยการเพิ่มขนาดหน่วยความจำของตัวโหนดที่เป็นเกตเวย์ให้มีขนาดมากขึ้น เพื่อเพิ่มการ

รองรับข้อมูลจากตัวโหนดเซนเซอร์ที่ต้องส่งเข้ามาที่ตัวโหนดที่เป็นเกตเวย์ ในขณะที่ไม่สามารถส่งข้อมูลออกนอกเครือข่ายได้ชั่วคราว แต่ก็ไม่สามารถแก้ปัญหาในกรณีที่โหนดเกิดความเสียหายจนไม่สามารถส่งและรับข้อมูลได้แบบถาวร

ดังนั้นงานวิจัยนี้จะแสดงให้เห็นถึงประโยชน์ และวิธีการในการเพิ่มโหนดที่เป็นเกตเวย์จาก 1 โหนด เป็น 2 โหนด ซึ่งในเบื้องต้นจะทำการเพิ่มจำนวนโหนดที่เป็นเกตเวย์แค่ 1 โหนด เนื่องจากสามารถแสดงให้เห็นถึงกรณีที่เกิด SPOF ที่โหนดได้ในทอพอโลยีแบบเดียวกัน ซึ่งโหนดที่เพิ่มจะอยู่ไกลกว่าโหนดเดิมและทดสอบโดยการส่งข้อมูลแบบ Constant Bit Rate (CBR) ผ่านโปรโตคอล Ad Hoc On-Demand Distance Vector (AODV) โดยจะทำการเปรียบเทียบปริมาณ Throughput ระหว่างโหนดที่เป็นเกตเวย์จาก 1 และ 2 โหนด ในกรณีที่โหนดที่เป็นเกตเวย์ 1 หยุดทำงาน



ภาพที่ 1: เครือข่ายเซนเซอร์ไร้สายแบบ 2 เกตเวย์

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

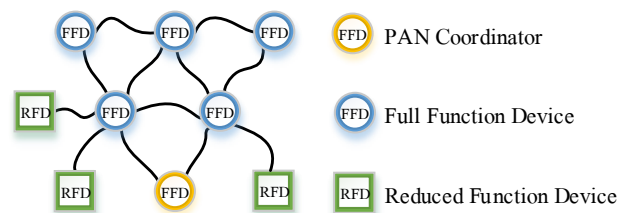
2.1 เครือข่ายเซนเซอร์ไร้สาย IEEE802.15.4

WSN มีหลายสถาบันที่ช่วยกันพัฒนา ซึ่ง Institute of Electrical and Electronics Engineers (IEEE) เป็นสถาบันที่มุ่งเน้นที่ชั้นที่ 1 กายภาพ (Physical layer) และชั้นที่ 2 ชั้นสื่อสารข้อมูล (Data link) ส่วน Internet Engineering Task Force (IETF) จะมุ่งเน้นไปที่ชั้นที่ 3 ชั้นเครือข่าย (Network layer) และชั้นที่สูงกว่าขึ้นไป ตามการจัดเรียงใน OSI model

โครงสร้างของ WSN จะมีโหนดที่เป็นเกตเวย์ ซึ่งทำหน้าที่ระบุช่องทางการสื่อสารภายในเครือข่ายระหว่างโหนดในเครือข่าย เพื่อไม่ทับซ้อนกัน และป้องกันการกวนกันของคลื่น

ความถี่ (Interference) โดยการสแกนหาระดับของพลังงานแต่ละช่องที่รับได้ ซึ่งโหนดใดๆ ที่อยู่ในช่องสัญญาณเดียวกันกับโหนดเกตเวย์ และต้องการจะเชื่อมต่อกับเครือข่ายเพื่อส่งข้อมูลไปยังตัวโหนดเกตเวย์ โหนดเหล่านั้นก็จะทำการส่งสัญญาณแจ้งเตือนออกมาพร้อมขอ (Beacon request) เพื่อขอเข้าร่วมเครือข่าย ซึ่งจะมีเพียงโหนดที่เป็นเกตเวย์เท่านั้นที่จะตอบสนองต่อการร้องขอเหล่านี้ได้

WSN จะประกอบไปด้วยโหนดต่างๆ จำนวนมาก แต่สามารถแบ่งออกเป็นสองแบบหลักคือ โหนดที่มีฟังก์ชันการทำงานแบบเต็ม Full-function device (FFD) และโหนดที่ถูกตัดฟังก์ชันที่ไม่จำเป็นออก Reduced-function device (RFD) ซึ่งในเครือข่ายจะต้องประกอบไปด้วย FFD อย่างน้อย 1 ตัว เพื่อทำหน้าที่เป็น PAN coordinator โดยตัวโหนดที่เป็น FFD จะสามารถทำหน้าที่เป็นโหนดเรเตอร์ และโหนดที่เป็นเกตเวย์ โดยโหนดเรเตอร์จะทำหน้าที่ในการจัดการเส้นทางให้กับโหนดต่างๆ ที่ต้องการส่งข้อมูลไปยังโหนดที่เป็นเกตเวย์ โดยโหนดนี้จะทำหน้าที่ในการจัดการเครือข่าย และสามารถทำหน้าที่เป็นเกตเวย์ โดยการเพิ่มส่วนของการติดต่อกับเครือข่ายอินเทอร์เน็ต เพื่อนำข้อมูลที่ได้อ่าน และโหนดประเภท RFD จะทำหน้าที่เป็นโหนดเซนเซอร์ ซึ่งจะเก็บข้อมูลจากตัวเซนเซอร์ต่างๆ ที่ทำงาน เช่น อุณหภูมิ จากนั้นจะส่งข้อมูลไปยังโหนดที่เป็นเกตเวย์ผ่านทางโหนดเรเตอร์



ภาพที่ 2: โครงสร้างของเครือข่ายเซนเซอร์ไร้สาย

2.2 Ad Hoc On-Demand Distance Vector (AODV)

AODV[4] เป็นโปรโตคอลเส้นทางตามมาตรฐาน RFC3561 ที่ใช้ทรัพยากรในการประมวลผล และหน่วยความจำน้อย จึงเหมาะกับ WSN ที่ต้องการเน้นเรื่องประหยัดพลังงาน AODV ใช้หลักการแบบ distance vector ของ Bellman-Ford ในการสร้างเส้นทาง และเป็นโปรโตคอลแบบ On-Demand ซึ่งจะทำ

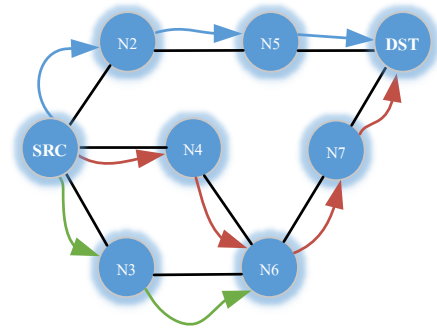
การสร้างเส้นทางเมื่อต้องการส่งข้อมูล ทำให้ลดการ Broadcast RREQ ได้ ตัวโปรโตคอล AODV จะพยายามทำให้โหนดนั้นๆ ได้รับเส้นทางโดยเร็วที่สุดเมื่อมีการกำหนดโหนดปลายทางใหม่ โดยอาศัยกระบวนการปรับปรุง และตรวจสอบเส้นทางที่ไปยังโหนดข้างเคียงว่าสามารถใช้งานได้หรือไม่ หรือเมื่อโครงสร้างเครือข่ายเปลี่ยนแปลง โดยให้โหนดนั้นๆ ยกเลิกการใช้เส้นทางที่ไม่สามารถใช้งานได้

โปรโตคอล AODV ประกอบไปด้วย Message ต่างๆ คือ Route Request (RREQ) Route Replies (RREP) Route Errors (RERR) เมื่อใดที่โหนดต้นทางต้องการส่งข้อมูลไปยังโหนดปลายทาง โหนดต้นทางจะ Broadcasts Message RREQ เพื่อค้นหาเส้นทางไปยังโหนดปลายทางนั้น และเส้นทางจะถูกสร้างขึ้นเมื่อ Message RREQ ไปถึงยังโหนดปลายทาง หรือในโหนดระหว่างเส้นทางนั้นมีเส้นทางที่สามารถไปยังโหนดปลายทางอยู่ในตารางเส้นทาง โหนดนั้นก็จะส่งส่ง Message RREP กลับไปยังโหนดต้นทางโดยวิธีการ Unicast

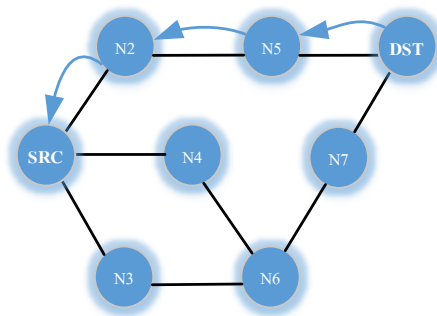
2.3 งานวิจัยที่เกี่ยวข้อง

งานวิจัยของ Xuan He, Yuanchen Ma และ Mika Mizutani ได้นำเสนอ Efficient Multiple Gateway System for WSN Management in BEMS [5] ซึ่งเป็นการเพิ่มโหนดโคออดิเนเตอร์ย่อยๆ ขึ้นมา แต่โหนดเหล่านี้ถ้าต้องการส่งข้อมูลออกนอกเครือข่าย จะต้องส่งผ่านโหนดที่เป็นเกตเวย์หลัก ดังนั้นเมื่อโหนดโคออดิเนเตอร์หลักได้รับความเสียหาย ทำให้เกิด Single of Failure (SPOF) ข้อมูลก็จะไม่สามารถส่งออกนอกเครือข่ายได้เช่นเดิม

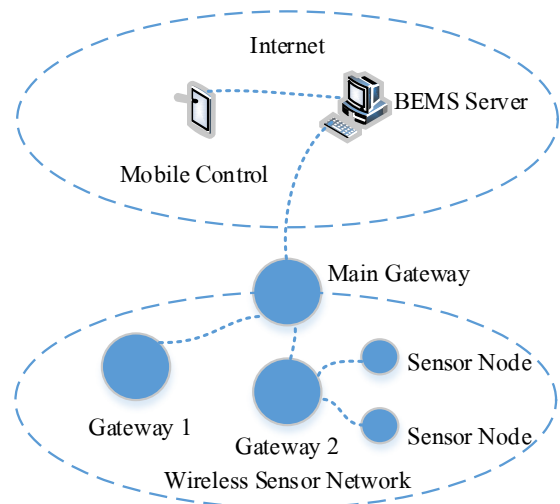
งานวิจัยของ Zahia Bidai, Hafid Haffaf, Moufida Maimour ได้นำเสนอ Node disjoint multi-path routing for zigbee cluster-tree Wireless Sensor Networks [6] โดยต้องการส่งข้อมูลที่ขนาดใหญ่ เช่น ภาพเคลื่อนไหว ผ่านทางหลายเส้นทางไปยังโหนดปลายทางโคออดิเนเตอร์เดียว ปัญหาที่เกิดขึ้นคือมีการชนกันของข้อมูลที่โหนดปลายทาง และมีการกวนกันของข้อมูลระหว่างเส้นทาง จึงไม่สามารถเพิ่มปริมาณ Throughput ได้มาก และไม่ส่งข้อมูลออกนอกเครือข่ายไปใช้งานได้ ในกรณีที่โหนดที่เป็นเกตเวย์เสีย จึงทำให้เกิด SPOF ในเครือข่ายได้



ภาพที่ 3: โหนดต้นทางกระจาย RREQ Message



ภาพที่ 4: โหนดปลายทางตอบ RREP Message



ภาพที่ 5: Efficient Multiple Gateway System for WSN Management in BEMS

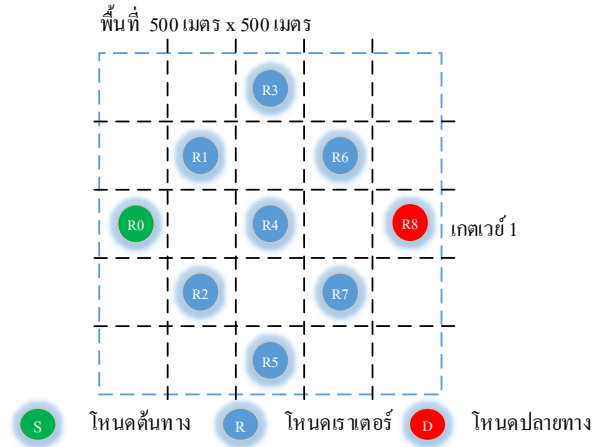
3. วิธีการดำเนินงาน

การจำลองการทำงานของระบบผ่านโปรแกรม Network Simulator (NS-2) โดยมีการสร้างทอพอโลยีที่ประกอบด้วยโหนดต้นทางจำนวน 1 โหนด และโหนดปลายทางจำนวน 1 และ 2 โหนด โดยมีการกำหนดช่วงเวลาการทำงานของเกตเวย์

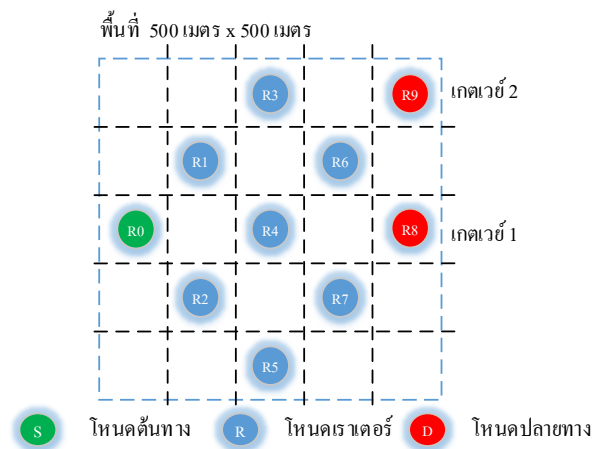
ดังภาพที่ 5 และทำการส่งข้อมูล CBR ที่มีจำนวนต่างๆ คือ 6000, 7000, 8000, 9000, 10000 packets เพื่อตรวจวัดค่า Throughput และค่า Packets Delivery Ratio (PDR) ของเครือข่าย ในกระบวนการจำลองการทำงานจะแบ่งเป็น 2 ส่วน

ส่วนที่ 1 จำลองการทำงานในกรณีที่มี 1 เกตเวย์ ทำงานเป็นระยะเวลา 50 วินาที โดยวินาทีที่ 5 จะเริ่มส่งข้อมูล CBR เข้าไปในโหนดต้นทาง R0 และกำหนดโหนดปลายทางเป็น R8 เมื่อเวลาผ่านไป 20 วินาที คือวินาทีที่ 25 กำหนดให้โหนดปลายทาง R8 หยุดทำงานจนกระทั่งถึงวินาทีที่ 45 รวมเป็นระยะเวลา 40 วินาที

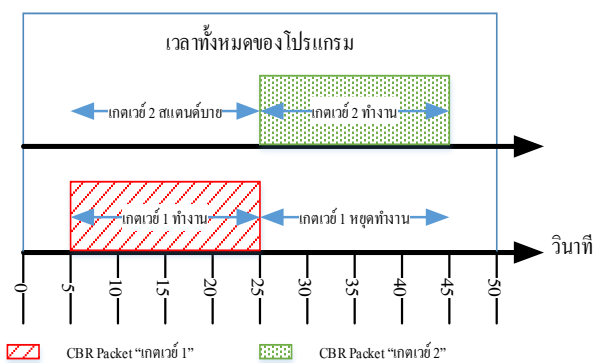
ส่วนที่ 2 จำลองการทำงานในกรณีที่มี 2 เกตเวย์ ทำงานเป็นระยะเวลา 50 วินาที โดยวินาทีที่ 5 จะเริ่มส่งข้อมูล CBR เข้าไปในโหนดต้นทาง R0 และกำหนดโหนดปลายทางเป็น R8 เมื่อเวลาผ่านไป 20 วินาที คือวินาทีที่ 25 กำหนดให้โหนดปลายทาง R8 หยุดทำงาน ในขณะนั้น โหนดที่เป็นเกตเวย์ 2 จะเริ่มทำงานแทนเมื่อรู้ว่าโหนดที่เป็นเกตเวย์ 1 หยุดทำงาน ซึ่งตัวโหนดที่เป็นเกตเวย์ 2 จะทำงานจนกระทั่งถึงวินาทีที่ 45 เป็นระยะเวลา 20 วินาที รวมเวลาในการจำลองทั้งหมด 40 วินาที ซึ่งแบ่งเป็นโหนดที่เป็นเกตเวย์ 1 ทำงาน 20 วินาที และโหนดที่เป็นเกตเวย์ 2 ทำงาน 20 วินาที



ภาพที่ 7: ทอพอโลยีแบบ 1 เกตเวย์



ภาพที่ 8: ทอพอโลยีแบบ 2 เกตเวย์



ภาพที่ 6: ช่วงเวลาการทำงานของเกตเวย์

การจำลองการทำงานจะกำหนดค่าพารามิเตอร์ในการทดสอบดังนี้

จำนวนโหนด (1 เกตเวย์)	9	โหนด
จำนวนโหนด (2 เกตเวย์)	10	โหนด
ชนิดข้อมูลที่ส่ง	Constant Bit Rate (CBR)	
รูปแบบการแพร่กระจายคลื่น	Two Ray Ground Model	
ขนาดข้อมูลที่ส่ง	100	ไบต์
ขนาดทอพอโลยี	500 เมตร x 500 เมตร	
โพรโทคอลระดับ MAC	IEEE 802.15.4	
โหนดต้นทาง	1	โหนด
โหนดปลายทาง (1 เกตเวย์)	1	โหนด
โหนดปลายทาง (2 เกตเวย์)	1	โหนด
ระยะเวลาในการจำลอง	40	วินาที
โพรโตคอลเส้นทาง	AODV	

หลังจากจำลองการทำงานของ WSN ทั้ง 2 แบบ แล้วจะได้ข้อมูลเป็น trace ไฟล์ ซึ่งบันทึกค่าเหตุการณ์ต่างๆ ไว้ที่เวลาต่างๆ โดยจะนำข้อมูลที่ได้อ่านนี้มาหาค่าเฉลี่ย Throughput และค่า Packets Delivery Ratio (PDR) ของเครือข่ายจากชุดโค้ด (Pseudo Code) ดังนี้

```

1 ALGORITHM THROUGHPUT (SOURCE_CBR, DESTINATION_CBR, TIME)
2 INPUT SOURCE_CBR # CBR_PACKET AT SOURCE NODE
3 INPUT DESTINATION_CBR # CBR_PACKET AT DESTINATION NODE
4 INPUT TIME # SIMULATION TIME
5 OUTPUT THROUGHPUT
6 THROUGHPUT_RATIO = 0
7 THROUGHPUT = ((DESTINATION_CBR/SOURCE_CBR)/TIME)*1000
8 RETURN (THROUGHPUT) #Throughput (Kbps)
9

```

ภาพที่ 9: การหาค่า Throughput

```

10 ALGORITHM PDR (SOURCE_CBR, DESTINATION_CBR)
11 INPUT SOURCE_CBR # CBR_PACKET AT SOURCE NODE
12 INPUT DESTINATION_CBR # CBR_PACKET AT DESTINATION NODE
13 OUTPUT PDR
14 PDR = 0
15 PDR = (DESTINATION_CBR/SOURCE_CBR)*100
16 RETURN (PDR) # Packet Delivery Ratio (PDR*)
17
18

```

ภาพที่ 10: การหาค่า PDR

ภาพที่ 9 แสดงการหาค่า Throughput ของ WSN แบบที่มี 1 เกตเวย์ โดยการหาค่า CBR packets ขาออกที่โหนด R8 เทียบกับขาเข้าที่โหนด R0 แล้วเทียบกับเวลาที่ใช้ในการจำลองคือ 40 วินาที ส่วน WSN แบบ 2 เกตเวย์ โดยการหาค่า CBR packets ขาออกที่โหนด R8 บวกกับโหนด R9 เทียบกับขาเข้าที่โหนด R0 แล้วเทียบกับเวลาที่ใช้ในการจำลองคือ 40 วินาที

ภาพที่ 10 แสดงการหาค่า PDR ของ WSN แบบ 1 เกตเวย์ โดยการหาค่า CBR packets ขาออกที่โหนด R8 เทียบกับขาเข้าที่โหนด R0 ส่วนของเครือข่ายแบบ 2 เกตเวย์ โดยการหาค่า CBR packets ขาออกที่โหนด R8 บวกกับโหนด R9 เทียบกับขาเข้าที่โหนด R0

4. ผลการทดลอง

4.1 การทดสอบการจำลองส่งข้อมูล CBR ใน WSN แบบ 1 เกตเวย์

ทำการทดสอบโดยการส่งข้อมูล CBR ที่มีอัตราข้อมูลที่แตกต่างกันเข้าไปที่โหนดต้นทาง R0 โดยมีโหนดปลายทาง R8

ตารางที่ 1: ผลการทดลองใน WSN แบบ 1 เกตเวย์

Parameter	Sending Data (Packets)				
	6000	7000	8000	9000	10000
Simulate time (s)	40	40	40	40	40
Data Receive R8 (Packets)	2999	3192	3192	3192	3192
Avg. Throughput R8 (Kbps)	12.50	11.40	9.98	8.87	7.98
Packet Delivery Ratio (R8) (%)	49.98	45.60	39.90	35.47	31.92
Packet Drops (R8) (Packets)	3001	3808	4808	5808	6808

จากตารางที่ 1 ค่า Throughput และค่า PDR จะเห็นได้ว่าเมื่อจำนวน CBR packets มีปริมาณมากขึ้นจะทำให้ Throughput และค่า PDR มีค่าที่น้อยลง เนื่องจากตัวโหนดไม่สามารถรองรับปริมาณข้อมูลที่มีจำนวนมากขึ้น ดังจะเห็นได้จากข้อมูลที่รับได้ในโหนด R8 ขณะที่จำนวน CBR packets ตั้งแต่ 7000 ขึ้นไปมีค่าเท่าเดิม และเมื่อพิจารณาในส่วนของจำนวน Packets Drops จะเห็นว่าจำนวนมากขึ้นเรื่อยๆ อันเป็นผลมาจากการที่โหนด R8 หยุดทำงาน

4.2 การทดสอบการจำลองส่งข้อมูล CBR ใน WSN แบบ 2 เกตเวย์

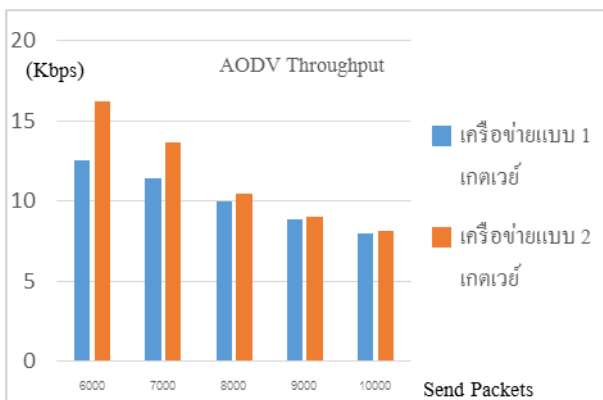
ทำการทดสอบโดยการส่งข้อมูล CBR เข้าไปที่โหนดต้นทาง R0 โดยมีโหนดปลายทาง R8 และ R9 จากตารางที่ 2 ผลที่ได้เมื่อมีการหาค่าปริมาณ Throughput และค่า PDR จะเห็นว่าที่จำนวนปริมาณ CBR packets ต่างๆ ปริมาณค่าของ Throughput และค่า PDR ของเครือข่ายที่มี 2 เกตเวย์ มีค่ามากกว่าเมื่อเทียบกับแบบเครือข่ายที่มี 1 เกตเวย์ เนื่องจากตัวโหนด R9 ที่ทำหน้าที่เป็นเกตเวย์ 2 สามารถช่วยรองรับปริมาณข้อมูลที่มีจำนวนมากขึ้นได้

จากกราฟในภาพที่ 11 และ 12 แสดงการเปรียบเทียบค่า Throughput และค่า PDR ของเครือข่ายแบบ 1 และ 2 เกตเวย์ ซึ่งจะเห็นได้ว่าเครือข่ายแบบ 2 เกตเวย์ สามารถทำได้สูงกว่าแบบเครือข่าย 1 เกตเวย์ ซึ่งปริมาณ Throughput ที่การส่ง Packets 6000 มีค่ามากกว่าที่ส่ง Packets 10000 นั้น เนื่องจากถ้าดูที่ตารางที่ 2 พบว่าจำนวน Packets Drop ของเกตเวย์ 2 มีค่า

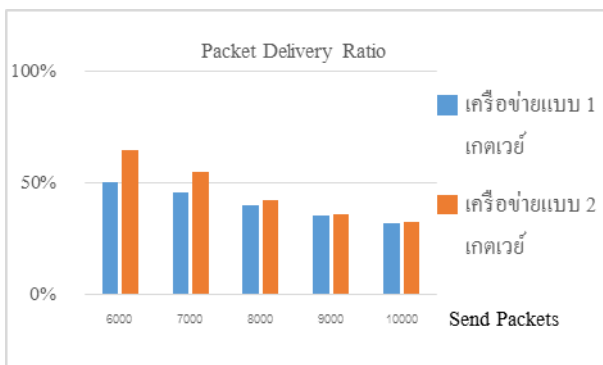
มากขึ้นเมื่อมีการส่งข้อมูลมากขึ้น เนื่องจากเกตเวย์ 2 อยู่ไกลสุด ทำให้ Buffer ที่โหนดระหว่างทางเต็ม ทำให้ Packets ที่ถูกส่งมาใหม่ถูก Drop ไป แต่โดยรวมของการจำลองที่ระดับ CBR packets ต่างๆ พบว่าในเครือข่าย 2 เกตเวย์ สามารถทำได้สูงกว่าแบบเครือข่าย 1 เกตเวย์

ตารางที่ 2: ผลการทดลองใน WSN แบบ 2 เกตเวย์

Parameter	Sending Data (Packets)				
	6000	7000	8000	9000	10000
Simulate time (s)	40	40	40	40	40
Data Received R8 (Packets)	2999	3192	3192	3192	3192
Data Received R9 (Packets)	887	634	159	54	54
Avg. Throughput R8+R9 (Kbps)	16.19	13.66	10.47	9.02	8.12
Packet Delivery Ratio R8+R9 (%)	64.77	54.66	41.89	36.07	32.46
Packet Drops R8+R9 (Packets)	2114	3174	4649	5754	6754



ภาพที่ 11: ปริมาณ Throughput



ภาพที่ 12: สัดส่วนการส่งข้อมูลสำเร็จ PDR

5. ข้อสรุปและเสนอแนะ

จากการทดลองแสดงให้เห็นว่า WSN แบบ 2 เกตเวย์ สามารถให้ปริมาณ Throughput ของเครือข่าย และค่า PDR% ที่ดีกว่าการใช้ WSN แบบ 1 เกตเวย์ โดยการใช้โปรโตคอลเส้นทาง AODV และโหนดภายในเครือข่ายไม่มีการเคลื่อนที่ แต่ในกรณีที่มีความหนาแน่นของข้อมูลในเครือข่ายจำนวนมาก ส่วนต่างของค่า Throughput ของเครือข่าย และค่า PDR นั้นจะมีค่าไม่สูงมาก โดยพบว่าค่า Throughput มีค่าสูงขึ้นเฉลี่ย 1.35 Kbps และค่า Packet Delivery Ratio (PDR) มีค่าสูงขึ้นเฉลี่ย 5.39% เมื่อเทียบกับ WSN แบบเดิมที่มี 1 เกตเวย์

ข้อเสนอแนะควรทำการปรับปรุงตัวโปรโตคอลเส้นทาง Ad-hoc On-Demand Distance Vector (AODV) ให้สามารถหาตัวเกตเวย์ใหม่ในเครือข่ายได้เองโดยอัตโนมัติ และแบ่งภาระการทำงานของตัวโหนดที่เป็นเกตเวย์ เมื่อเกตเวย์ปัจจุบันไม่สามารถส่งหรือรับข้อมูลได้

เอกสารอ้างอิง

- [1] อนุรักษ์ เขยงมุ่ม และ ชัยพร เขมระภาคะพันธ์, "การประยุกต์เครือข่ายตรวจจับสัญญาณไร้สายเพื่อประมวลผลข้อมูลความเร็วลมสำหรับช่วยตัดสินใจในการติดตั้งกังหันลมผลิตกระแสไฟฟ้า", *การประชุมการวิจัยระดับบัณฑิตศึกษาแห่งชาติ มหาวิทยาลัยเชียงใหม่ ครั้งที่ 2*, 26 พฤศจิกายน 2553, หน้า 620-629.
- [2] นภาพร พิมปรู และ ประโยชน์ คำสวัสดิ์, "การบีบอัดข้อมูลในเครือข่ายเซนเซอร์ไร้สาย" *ECTI-CARD 2013* หน้า 173-178.
- [3] Luca De Nardis and Maria-Gabriella, "Overview of the IEEE 802.15.4/4a standards for low data rate Wireless Personal Data Networks" *WPNC 2007*, Handover, Germany, March 2007.
- [4] P. Manoyut and W. Suntiamorntut, "Modified Local Repair in AODV for LoWPANs" (*IMCOM 2012*), Kuala Lumpur, Malaysia, February 2012.
- [5] Xuan He and Yuanchen Ma, "Efficient Multiple Gateway System for WSN Management in BEMS" *Networked Sensing Systems (INSS)*, Belgium, June 2012.
- [6] Zahia Bidai and Hafid Haffaf "Node disjoint multi-path routing for zigbee cluster-tree Wireless Sensor Networks", *Multimedia Computing and Systems (ICMCS)*, Ouarzazate, Morocco, April 2011.

การศึกษาปัจจัยที่ส่งผลต่อการรับสัญญาณในการสื่อสารไร้สายภายใต้สถานะแวดล้อมต่างๆ

A Study of Factors Effecting Wireless Networking in any Environments

นายสุรพล ไกรจักร (Surapon Kraijak)¹ และนายปานวิทย์ ชูวะนุติ (Panwit Tuwanuti)²

¹ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ สถาบันพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

²ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ สถาบันพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

macro_biggy@hotmail.com, panwit@it.kmitl.ac.th

บทคัดย่อ

บทความนี้นำเสนอการศึกษาปัจจัยที่ส่งผลต่อการรับสัญญาณในการสื่อสารไร้สายภายใต้สถานะแวดล้อมต่างๆ โดยมุ่งเน้นการค้นคว้าและวิจัยแบบจำลองการลดทอนตามระยะทางและแบบจำลองการจางหายของกำลังสัญญาณในแต่ละสถานะแวดล้อมที่ได้คิดค้นขึ้นตั้งแต่ในอดีตจนถึงปัจจุบัน ซึ่งในบทความฉบับนี้จะแบ่งการทดลองออกเป็น 2 รูปแบบ คือ การทดลองโดยการเลียนแบบจากโปรแกรมคอมพิวเตอร์ และการทดลองในสถานะแวดล้อมจริง เพื่อนำผลการทดลองที่ได้จากการทดลองทั้ง 2 รูปแบบในแต่ละสถานะแวดล้อมมาเปรียบเทียบเพื่อหาว่าแบบจำลองการลดทอนตามระยะทางและแบบจำลองการจางหายของกำลังสัญญาณชนิดใดที่ก่อให้เกิดความแม่นยำ ถูกต้อง และเหมาะสมสำหรับการนำไปพยากรณ์อัตราการลดทอนตามระยะทางและอัตราการจางหายของกำลังสัญญาณในสถานะแวดล้อมนั้นมากที่สุด เพื่อนำไปใช้เป็นข้อมูลประกอบการออกแบบระบบโครงข่ายไร้สายต่อไป

คำสำคัญ: การติดต่อสื่อสารไร้สาย การลดทอนตามระยะทาง การจางหายของกำลังสัญญาณ

Abstract

This paper presents a study of factors effecting wireless network in any environment which focus on investigated and researched of path loss and fading model which designed in a past until now in any environments. This paper divided the experiment into 2 systems are an experiment by using computer simulation and experiment in real environments to compared the experimental results from both systems and found out the accurate, correct and suitable path loss and fading model in each environments that useful for a designing of wireless network communication.

Keyword: Wireless communication, Path loss, Fading.

1. บทนำ

ในปัจจุบันเทคโนโลยีการสื่อสารแบบไร้สายได้เข้ามามีบทบาทสำคัญในการดำเนินกระบวนการทางธุรกิจขององค์กรอย่างหลีกเลี่ยงไม่ได้ ทำให้ผู้บริหารเริ่มมีความตระหนักถึงการนำเทคโนโลยีดังกล่าวมาใช้ภายในองค์กรมากขึ้น แต่ด้วยข้อจำกัดที่ผู้บริหารองค์กรส่วนใหญ่ มักขาดความรู้และทักษะทางด้านเทคโนโลยีสารสนเทศ จึงส่งผลทำให้เกิดการวิเคราะห์ ออกแบบ และนำเทคโนโลยีการสื่อสารแบบไร้สายไปใช้อย่างไม่เหมาะสม ในเวลาต่อมาองค์กรธุรกิจได้เริ่มมีการจัดตั้งแผนกเทคโนโลยีสารสนเทศขึ้น ซึ่งแผนกดังกล่าวจะทำหน้าที่ในการวิเคราะห์ ออกแบบ แก้ไขปัญหา และดูแลรักษาระบบเครือข่ายขององค์กรให้สามารถทำงานได้อย่างมีประสิทธิภาพและปราศจากข้อบกพร่อง แต่การจะวิเคราะห์ ออกแบบเครือข่ายให้มีประสิทธิภาพสูงที่สุดนั้นจำเป็นต้องอาศัยแบบจำลองที่มีความเหมาะสมและให้ผลลัพธ์ที่มีความใกล้เคียงกับสถานะแวดล้อมจริงมากที่สุด ซึ่งเป็นเรื่องที่ยากสำหรับนักคอมพิวเตอร์ในการเลือกแบบจำลองมาใช้ในการทดลองหรือวิเคราะห์ออกแบบ เนื่องจากแบบจำลองที่ถูกคิดค้นขึ้นในปัจจุบันมีอยู่มากมาย และไม่สามารถทราบได้เลยว่าแบบจำลองใดมีความเหมาะสมกับสถานะแวดล้อมแบบใด ซึ่งถ้าหากเลือกแบบจำลองผิดพลาดก็อาจจะส่งผลต่อการวิเคราะห์ ออกแบบ รวมไปถึงประสิทธิภาพการทำงานของระบบเครือข่ายในระยะยาวได้

โดยแนวความคิดในการแก้ปัญหาที่ถูกคิดค้นขึ้นในอดีต จะมีอยู่หลายรูปแบบด้วยกัน แต่ก็ยังไม่สามารถนำไปใช้เป็นแนวทางในการแก้ปัญหาที่เกิดขึ้นได้อย่างมีประสิทธิภาพ เช่นในงานวิจัยของ W. Afric, B. Z. Cihlar และ Sonja Grgic [1] ในปี 2007 ได้กล่าวถึง แนวทางในการคำนวณหาการลดทอนตามระยะทางโดยใช้ผลที่ได้จากการตรวจวัด ซึ่งข้อจำกัดของ

งานวิจัยดังกล่าวคือ ขาดความหลากหลายในการทดลองที่สถานะแวดล้อมอื่นๆ จึงส่งผลให้ผลการทดลองที่ได้ นำไปใช้อ้างอิงได้เพียงในบางสถานะแวดล้อมเท่านั้น และยังมีข้อจำกัดในการนำผลการทดลองที่ได้ไปใช้ในการอ้างอิงอีกด้วย เนื่องจากผลการทดลองที่ปรากฏในบทความวิจัยฉบับนี้เป็นการคำนวณการเกิดการลดทอนตามระยะทางบนเครือข่ายไวแมกซ์ (WiMAX) ในย่านความถี่ 3.5 GHz. เท่านั้น ในงานวิจัยของ Mohammed Alshami, Tugral Arslan, J. Thompson และ Armet T. Erdogan [2] ในปี 2011 ได้กล่าวถึง การวิเคราะห์แบบการลดทอนตามระยะทางเชิงความถี่บนเทคโนโลยีการสื่อสารไวแมกซ์ ซึ่งข้อจำกัดของงานวิจัยดังกล่าวคือ ขาดการอธิบายให้เห็นถึงความน่าเชื่อถือของผลการทดลองการลดทอนตามระยะทางในสถานะแวดล้อมจริงที่นำมาใช้ภายในบทความ และการเขียนสรุปผลการทดลองไม่ได้แสดงให้เห็นถึงสัดส่วนความคลาดเคลื่อนระหว่างผลลัพธ์ที่ได้จากการทดลองโดยการเขียนแบบจากโปรแกรมคอมพิวเตอร์กับผลลัพธ์ที่ได้จากการทดลองในสถานะแวดล้อมจริงอย่างชัดเจน

จากข้อจำกัดที่ได้กล่าวในข้างต้น จึงก่อเกิดแนวความคิดในการศึกษาและเปรียบเทียบแบบจำลองที่เกี่ยวข้องกับการวิเคราะห์และออกแบบระบบเครือข่ายไร้สาย เช่น แบบจำลองการลดทอนตามระยะทาง แบบจำลองการจางหายของกำลังสัญญาณ ที่มีความหลากหลายมากยิ่งขึ้น เพื่อให้ให้นักวิเคราะห์และออกแบบระบบเครือข่ายไร้สาย หรือผู้ที่สนใจสามารถนำไปใช้เป็นแนวทางประกอบการวิเคราะห์ ออกแบบหรือนำไปต่อยอดในงานวิจัยของตนเองต่อไป ตลอดจนเพื่อลดเวลาและข้อผิดพลาดในการวิเคราะห์ออกแบบระบบเครือข่ายไร้สาย อันจะส่งผลดีต่อองค์กร โดยรวมในแง่ของการลดต้นทุนและเพิ่มประสิทธิภาพ รวมทั้งผลิตภาพ (Productivity) ในการทำงานของพนักงานภายในองค์กรให้เพิ่มมากขึ้นด้วย

2. ทฤษฎีที่เกี่ยวข้อง

ในหัวข้อนี้จะอธิบายถึงทฤษฎีที่นำมาใช้ประกอบการศึกษาค้นคว้าและวิจัยภายในบทความ ซึ่งมีรายละเอียดดังนี้

2.1 ทฤษฎีแบบจำลองการลดทอนตามระยะทาง

การลดทอนตามระยะทาง (Path loss) เป็นปรากฏการณ์ที่เกิดขึ้นเมื่อสัญญาณที่ได้รับอ่อนลง เนื่องจากระยะห่างระหว่างเครื่องรับและเครื่องส่งที่เพิ่มมากขึ้น สำหรับในกรณีที่เครื่องรับ

อยู่ไกลจากเครื่องส่ง สัญญาณที่ถูกส่งออกไปจะมีการสูญเสียมากขึ้นตามกำลังสองของระยะทางที่เพิ่มขึ้นและกำลังสองของความถี่ที่ใช้ด้วย โดยทั่วไปแบบจำลองการลดทอนตามระยะทางสามารถจำแนกตามความเหมาะสมสำหรับการนำไปใช้ทดลองในสถานะแวดล้อมต่างๆ ได้ 4 แบบ ดังนี้

2.1.1 สถานะแวดล้อมแบบพื้นที่เปิดโล่ง

แบบจำลองการลดทอนตามระยะทางในสถานะแวดล้อมแบบพื้นที่เปิดโล่งที่นำมาทดลองภายในบทความฉบับนี้มีอยู่ทั้งหมด 3 แบบจำลองด้วยกัน ดังนี้ [2]

2.1.1.1 Simplified model

$$PL_{fs} = 32.45 + 20\log_{10}(d) + 20\log_{10}(f)$$

2.1.1.2 Walfish model

$$L_{fs} = 32.45 + 20\log_{10}(d) + 20\log_{10}(f) + s$$

2.1.1.3 Lee model

$$PL = 96.92 + 20\log_{10}(d) + 10\log_{10}(f/900) - \alpha_0$$

2.1.2 สถานะแวดล้อมแบบชุมชนเมือง

แบบจำลองการลดทอนตามระยะทางในสถานะแวดล้อมแบบชุมชนเมืองที่นำมาทดลองภายในบทความฉบับนี้มีอยู่ทั้งหมด 7 แบบจำลองด้วยกัน ดังนี้ [2] [3]

2.1.2.1 Ericsson model

$$PL = 36.2 + 30.2\log_{10}(d) + 12\log_{10}(h_b) + 0.1\log_{10}(h_b) * \log_{10}(d) - 3.2(\log_{10}(11.75 * h_m))^2 + 44.49\log_{10}(f) - (4.78\log_{10}(f))^2$$

2.1.2.2 Hata model

$$PL = 69.55 + 26.16\log_{10}(f) - 13.82\log_{10}(h_b) + (44.9 - 6.55\log_{10}(h_b)) * \log_{10}(d) + s (1.11\log_{10}(f) - 0.7) * h_m (1.56\log_{10}(f) - 0.8)$$

2.1.2.3 COST-231 model

$$PL = 43.30 + 33.90\log_{10}(f) - 13.82\log_{10}(h_b) + (44.9 - 6.55\log_{10}(h_b)) * \log_{10}(d) + s (1.11\log_{10}(f) - 0.7) * h_m (1.56\log_{10}(f) - 0.8)$$

2.1.2.4 Erceg model

$$PL = 20 * \log_{10}(4 * 3.14 * d_0 * (f / C)) + 10 * (4.6 - 0.0075 * h_b + 12.6 / h_b) * \log_{10}(d / d_0) + S + 6\log_{10}(f / 2000) - 10.8\log_{10}(h_m / 2000)$$

2.1.2.5 ECC-33 model

$$PL = 92.4 + 20\log_{10}(d) + 20\log_{10}(f) + 20.41 + 9.83\log_{10}(d) + 7.897\log_{10}(f) + 9.56(\log_{10}(f))^2 - \log_{10}(h_b/2000)\{13.98 + 5.8[\log_{10}(d)^2]\} - (42.57 + 13.7\log_{10}(f))[\log_{10}(h_m) - 0.585]$$

2.1.2.6 Lee model

$$PL = 123.77 + 30.5\log_{10}(d) + 10\log_{10}(f/900) - \alpha_0$$

2.1.2.7 COST231 Walfish-Ikegami model

$$PL = 49.186 + 38\log_{10}(d) + (25.3 + 0.7 * (f/925)) * \log_{10}(f) + 6\log_{10}(f/2000)$$

2.1.3 สถานะแวดล้อมแบบชนเมือง

แบบจำลองการลดทอนตามระยะทางในสถานะแวดล้อมแบบชนเมืองที่นำมาทดลองภายในบทความฉบับนี้มีอยู่ทั้งหมด 8 แบบจำลองด้วยกัน ดังนี้ [2] [3] [4]

2.1.3.1 Okumura model

$$PL = 92.4 + 20\log_{10}(d) + 20\log_{10}(f) + A_{mm}(f, d) - (20\log_{10}(h_b/2000)) - (20\log_{10}(h_m/3)) - 28$$

2.1.3.2 Ericsson model

$$PL = 43.20 + 68.93\log_{10}(d) + 12\log_{10}(h_b) + 0.1\log_{10}(h_b) * \log_{10}(d) - 3.2(\log_{10}(11.75 * h_m))^2 + 44.49\log_{10}(f) - (4.78\log_{10}(f))^2$$

2.1.3.3 Hata model

$$PL = 69.55 + 26.16\log_{10}(f) - 13.82\log_{10}(h_b) + (44.9 - 6.55\log_{10}(h_b))\log_{10}(d)\log_{10}(d) + s(1.11\log_{10}(f) - 0.7) * h_m(1.56\log_{10}(f) - 2(\log_{10}(f/28))^2) - 5.4$$

2.1.3.4 COST-231 model

$$PL = 43.30 + 33.90\log_{10}(f) - 13.82\log_{10}(h_b) + (44.9 - 6.55\log_{10}(h_b)) * \log_{10}(d) + s(1.11\log_{10}(f) - 0.7) * h_m(1.56\log_{10}(f) - 0.8 - 2(\log_{10}(f/28))^2) - 5.4$$

2.1.3.5 Lee model

$$PL = 99.86 + 38.4\log_{10}(d) + 10\log_{10}(f/900) - \alpha_0$$

2.1.3.6 Erceg model

$$PL = 20 * \log_{10}(4 * 3.14 * d_0 * (f/C)) + 10 * (4.0 - 0.0065 * h_b + 17.1/h_b) * \log_{10}(d/d_0) + S + 6\log_{10}(f/2000) - 10.8\log_{10}(h_m/2000)$$

2.1.3.7 ECC-33 model

รูปแบบของสมการเหมือนกับสมการที่ 2.1.2.5

2.1.3.8 Okumura-Hata model

$$PL = 20\log_{10}(4 * 3.14 * d_0 / C) + 10 * (a - b * h_b + c / h_b) * \log_{10}(d/d_0) + s, d > d_0$$

2.1.4 สถานะแวดล้อมแบบชนบท

แบบจำลองการลดทอนตามระยะทางในสถานะแวดล้อมแบบชนบทที่นำมาทดลองภายในบทความฉบับนี้มีอยู่ทั้งหมด 6 แบบจำลองด้วยกัน ดังนี้ [2] [4] [5]

2.1.4.1 Okumura model

รูปแบบของสมการเหมือนกับสมการที่ 2.1.3.1

2.1.4.2 Ericsson model

$$PL = 45.95 + 100.6\log_{10}(d) + 12\log_{10}(h_b) + 0.1 * \log_{10}(h_b) * \log_{10}(d) - 3.2(\log_{10}(11.75 * h_m))^2 + 44.49\log_{10}(f) - (4.78\log_{10}(f))^2$$

2.1.4.3 Hata model

$$PL = 69.55 + 26.16\log_{10}(f) - 13.82\log_{10}(h_b) + (44.9 - 6.55\log_{10}(h_b))\log_{10}(d)\log_{10}(d) + s(1.11\log_{10}(f) - 0.7) * h_m(4.78(\log_{10}(f))^2) + (18.33\log_{10}(f) - 40.98)$$

2.1.4.4 Erceg model

$$PL = 20 * \log_{10}(4 * 3.14 * d_0 * (f/C)) + 10 * (3.6 - 0.005 * h_b + 20/h_b) * \log_{10}(d/d_0) + S + 6\log_{10}(f/2000) - 10.8\log_{10}(h_m/2000)$$

2.1.4.5 COST-231 model

$$PL = 43.30 + 33.90\log_{10}(f) - 13.82\log_{10}(h_b) + (44.9 - 6.55\log_{10}(h_b)) * \log_{10}(d) + s(1.11\log_{10}(f) - 0.7) * h_m(1.56\log_{10}(f) - 0.8 - 4.78(\log_{10}(f))^2) - 40.98 + 18.33\log_{10}(f)$$

2.1.4.6 Lee model

$$PL = 86.12 + 43.5\log_{10}(d) + 10\log_{10}(f/900) - \alpha_0$$

2.2 ทฤษฎีแบบจำลองการจางหายของกำลังสัญญาณ

การจางหายของกำลังสัญญาณ (Fading) หมายถึง การที่สัญญาณถูกลดทอนลงไป ทำให้มีความแรงของสัญญาณอ่อนลง โดยการจางหายของกำลังสัญญาณสามารถจำแนกออกได้เป็น 2 ประเภทด้วยกันคือ แพลตเฟดดิ้ง (Flat Fading) และ ฟริควเอนซีซีเล็กเฟดดิ้งเฟดดิ้ง (Frequency Select Fading) ซึ่งการจางหายของกำลังสัญญาณทั้งสองประเภทนั้น ไม่สามารถที่จะพยากรณ์การจางหายของกำลังสัญญาณออกเป็นตัวเลขที่แน่นอนได้ เนื่องจากปริมาณการจางหายของกำลังสัญญาณจะขึ้นอยู่กับคุณสมบัติของชั้นบรรยากาศ ณ ขณะนั้น สภาพอากาศ และสภาพภูมิประเทศในแต่ละพื้นที่ ซึ่งจะก่อให้เกิดการจาง

หายของกำลังสัญญาณในปริมาณที่ไม่เท่ากัน จากปัญหาดังกล่าวทำให้นักวิจัยทางด้านโทรคมนาคมได้ทำการคิดค้นแบบจำลองการจางหายของกำลังสัญญาณที่มีความน่าเชื่อถือขึ้นหลายแบบจำลองด้วยกัน โดยแบบจำลองที่ได้ทำการคัดเลือกและนำมาใช้ในการทดลองในบทความฉบับนี้มีอยู่ทั้งหมด 5 แบบจำลองด้วยกัน ดังนี้ [4] [5]

2.2.1 Ricean distribution

$$P(r) = (r / \sigma) \exp[-(r^2 + A^2) / (2 * (\sigma)^2)] * I_0 * (A_r / (\sigma)^2)$$

2.2.2 Rayleigh distribution

$$P(r) = (r / (\sigma)^2) \exp[-r^2 / (2 * (\sigma)^2)]$$

2.2.3 Log-Normal Fading model

$$P(r) = (1 / r \sqrt{2 * 3.14 * (\sigma)^2}) \exp[-(\ln(r) - m)^2 / (2(\sigma)^2)]$$

2.2.4 Nakagami model

$$P(r) = (2m^m r^{(2m-1)} * \exp[-mr^2 / \omega]) / \Gamma(m) * \omega^m$$

2.2.5 Weibull model

$$P(r) = (\alpha * b / r_0) * \exp[-(br / r_0)^\alpha] * (br / r_0)^{\alpha-1}$$

โดยแบบจำลองข้างต้นถือเป็นแบบจำลองที่มีความแม่นยำสูงและถูกนำไปใช้กันอย่างแพร่หลายสำหรับการคาดการณ์อัตราการจางหายของกำลังสัญญาณในสภาวะแวดล้อมใดๆ

ตารางที่ 1: แสดงและกำหนดค่าพารามิเตอร์ที่เกี่ยวข้องกับการทดลอง

ชื่อตัวแปร	คำอธิบาย	ค่าพารามิเตอร์
f	ความถี่ ในหน่วย MHz.	2400
D	ระยะทาง ในหน่วยกิโลเมตร	5
D_0	ระยะทางอ้างอิง ในหน่วยเมตร	100
C	ขนาดความยาวคลื่น	0.125*
S	ขอบเขตการเกิดรบกวน ในหน่วย dB	10
H_b	ความสูงสายอากาศฝั่งส่ง ในหน่วยเมตร	5
H_m	ความสูงสายอากาศฝั่งรับ ในหน่วยเมตร	5
G_b	อัตราขยายของสายอากาศฝั่งส่ง (dB)	10
G_m	อัตราขยายของสายอากาศฝั่งรับ (dB)	10
s	ส่วนเบี่ยงเบนมาตรฐาน ในหน่วย dB	8.2-10.6
r	ค่าแอมพลิจูดของสัญญาณที่รับได้	0 – 6 Volt.
A	ค่าแอมพลิจูดสูงสุดของสัญญาณที่รับได้	2 Volt.

*ผลที่ได้จากการคำนวณด้วยสูตร $C = 3 * 10^8 / f$

3. การวิเคราะห์และออกแบบการทดลอง

การวิเคราะห์และออกแบบการทดลอง ถือเป็นกระบวนการที่มีความสำคัญอย่างมากสำหรับการจัดทำโครงการวิจัย ดังนั้นหัวข้อนี้จึงเป็นการกล่าวถึงรูปแบบและขั้นตอนการทดลองเป็นสำคัญ โดยในบทความฉบับนี้ได้ทำการจำแนกการทดลองออกเป็น 2 รูปแบบ คือ การทดลองโดยการเลียนแบบจากโปรแกรมคอมพิวเตอร์ และการทดลองในสภาวะแวดล้อมจริง ซึ่งมีรายละเอียดดังนี้

3.1 การทดลองโดยการเลียนแบบจากโปรแกรมคอมพิวเตอร์

ในการทดลองโดยการเลียนแบบจากโปรแกรมคอมพิวเตอร์ภายในบทความฉบับนี้จะใช้โปรแกรม MATLAB สำหรับการเลียนแบบตามแบบจำลองการลดทอนตามระยะทางและการจางหายของกำลังสัญญาณที่ได้นำเสนอในหัวข้อ 2.1 และ 2.2 ตามลำดับ โดยการทดลองในรูปแบบนี้จะทำการทดลองในสภาวะแวดล้อมที่แตกต่างกัน 4 แบบ คือ สภาวะแวดล้อมแบบพื้นที่เปิดโล่ง สภาวะแวดล้อมแบบชุมชนเมือง สภาวะแวดล้อมแบบชานเมือง และสภาวะแวดล้อมแบบชนบท ซึ่งขั้นตอนการทดลองในสภาวะแวดล้อมต่างๆ จะประกอบด้วย ขั้นตอนแรกทำการสร้างแบบจำลองการลดทอนตามระยะทางและการจางหายของกำลังสัญญาณลงบนโปรแกรม MATLAB ขั้นตอนที่ 2 ทำการประมวลผลการเลียนแบบตามแบบจำลองการลดทอนตามระยะทางและแบบจำลองการจางหายของกำลังสัญญาณในสภาวะแวดล้อมต่างๆ ขั้นตอนที่ 3 บันทึกผลการทดลองและนำผลลัพธ์ดังกล่าวไปบันทึกลงบนกราฟ ขั้นตอนที่ 4 สรุปและวิเคราะห์ผลการทดลอง

3.2 การทดลองในสภาวะแวดล้อมจริง

การทดลองในสภาวะแวดล้อมจริง จำเป็นต้องอาศัยอุปกรณ์กำเนิดและรับสัญญาณคลื่นวิทยุในย่านความถี่ 2.4 GHz. ในการทดลองการลดทอนตามระยะทางและการจางหายของกำลังสัญญาณในสภาวะแวดล้อมทั้ง 4 แบบ เช่นเดียวกันกับการทดลองโดยการเลียนแบบจากโปรแกรมคอมพิวเตอร์ โดยอุปกรณ์ที่ใช้ทดลองในสภาวะแวดล้อมจริงในบทความฉบับนี้มีทั้งหมด 2 ชิ้นด้วยกันคือ เครื่องกำเนิดสัญญาณ IFR2025 และเครื่องรับสัญญาณ HP8560A ซึ่งขั้นตอนการทดลองในสภาวะแวดล้อมจริง จะประกอบด้วย ขั้นตอนแรก กำหนดพื้นที่ในการทดลองจริงตามแต่ละสภาวะแวดล้อม ซึ่งสามารถจำแนกได้ดังนี้ การทดลองในสภาวะแวดล้อมแบบพื้นที่เปิดโล่ง ใช้พื้นที่

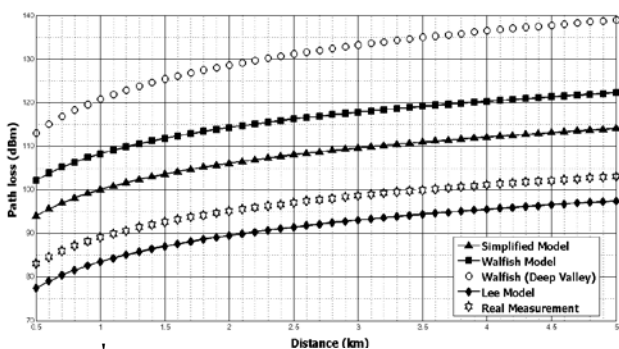
บริเวณถนนเลียบสนามบินสุวรรณภูมิ เขตลาดกระบัง จังหวัด กรุงเทพมหานคร การทดลองในสภาวะแวดล้อมแบบชุมชนเมือง ใช้พื้นที่บริเวณแฟลตการเคหะร่มเกล้า เขตลาดกระบัง จังหวัดกรุงเทพมหานคร การทดลองในสภาวะแวดล้อมแบบชานเมือง ใช้พื้นที่บริเวณหน้าโชว์รูมโตโยต้า (หัวตะเข้) เขตลาดกระบัง จังหวัดกรุงเทพมหานคร และการทดลองในสภาวะแวดล้อมแบบชนบท ใช้พื้นที่บริเวณซอยพัฒนาชนบท 3 เขตลาดกระบัง จังหวัดกรุงเทพมหานคร โดยในแต่ละสภาวะแวดล้อมจะทำการทดลองเป็นระยะทางประมาณ 5 กิโลเมตร ขั้นตอนที่ 2 ติดตั้งเครื่องกำเนิดและเครื่องรับสัญญาณคลื่นวิทยุตามจุดทดลองต่างๆ ขั้นตอนที่ 3 วัดระดับสัญญาณบนเครื่องรับที่ได้รับ ได้ตามระยะทางที่กำหนด ขั้นตอนที่ 4 บันทึกผลการทดลองและนำผลลัพธ์ดังกล่าวไปบันทึกลงบนกราฟ ขั้นตอนที่ 5 สรุปผลและวิเคราะห์ผลการทดลองเปรียบเทียบกับผลการทดลองโดยการเลียนแบบจากโปรแกรมคอมพิวเตอร์

4. การทดลองและผลการทดลอง

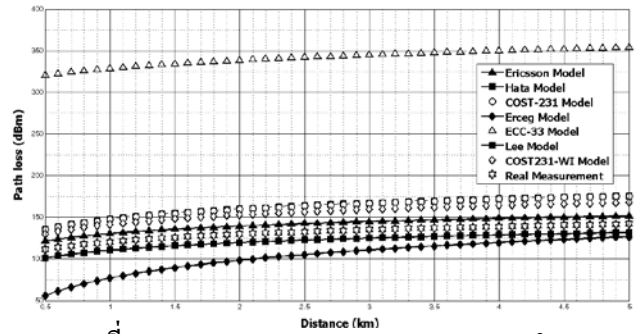
ในหัวข้อนี้จะกล่าวถึงการทดลองและผลการทดลองที่ได้จากการเลียนแบบโดยโปรแกรมคอมพิวเตอร์และการทดลองในสภาวะแวดล้อมจริงในแต่ละสภาวะแวดล้อมที่กำหนด โดยจะจำแนกการนำเสนอผลการทดลองออกเป็น 2 ส่วนด้วยกันคือ ผลการทดลองการลดทอนตามระยะทาง และผลการทดลองการจางหายของกำลังสัญญาณ ซึ่งมีรายละเอียดดังนี้

4.1 ผลการทดลองการลดทอนตามระยะทาง

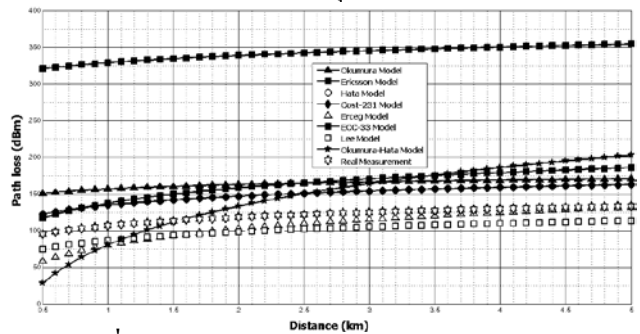
ในส่วนนี้จะเป็นการนำเสนอผลการทดลองการลดทอนตามระยะทางที่ได้จากการเลียนแบบโดยโปรแกรมคอมพิวเตอร์และการทดลองในสภาวะแวดล้อมจริงในแต่ละสภาวะแวดล้อมที่กำหนด โดยแกนนอนคือระยะทางและแกนตั้งคือระดับสัญญาณที่ได้รับ ซึ่งมีรายละเอียดดังนี้



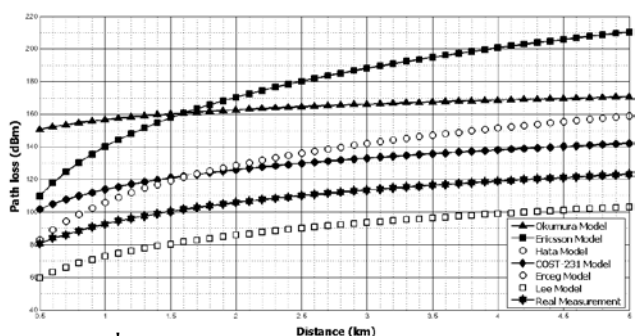
ภาพที่ 1: แสดงผลการทดลองการลดทอนตามระยะทางในสภาวะแวดล้อมแบบพื้นที่เปิดโล่ง



ภาพที่ 2: แสดงผลการทดลองการลดทอนตามระยะทางในสภาวะแวดล้อมแบบชุมชนเมือง



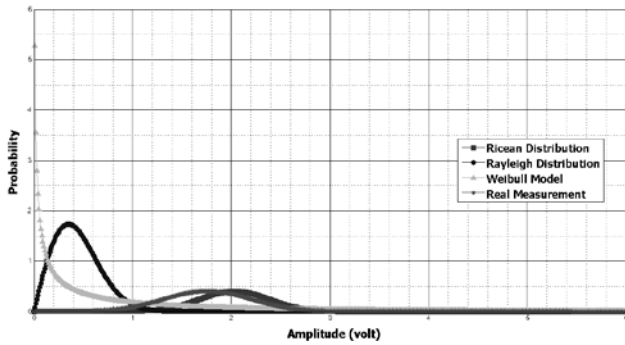
ภาพที่ 3: แสดงผลการทดลองการลดทอนตามระยะทางในสภาวะแวดล้อมแบบชานเมือง



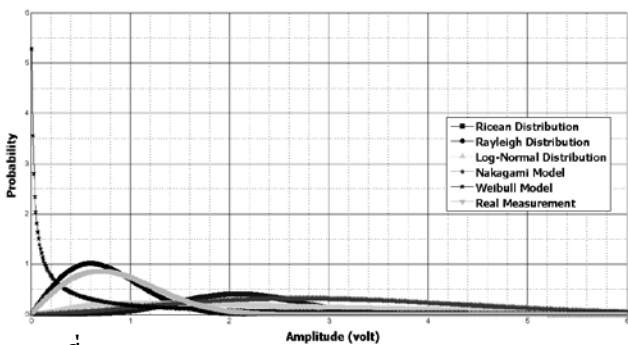
ภาพที่ 4: แสดงผลการทดลองการลดทอนตามระยะทางในสภาวะแวดล้อมแบบชนบท

4.2 ผลการทดลองการจางหายของกำลังสัญญาณ

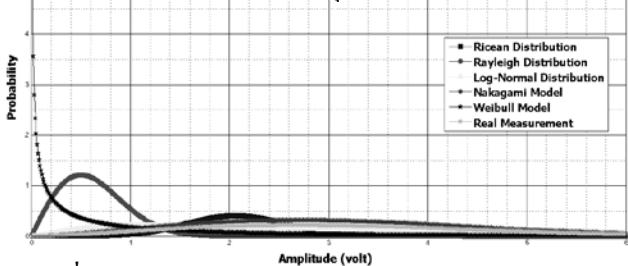
ในส่วนนี้จะเป็นการนำเสนอผลการทดลองการจางหายของกำลังสัญญาณที่ได้จากการทดลองในสภาวะแวดล้อมจริงและการเลียนแบบโดยโปรแกรมคอมพิวเตอร์ในแต่ละสภาวะแวดล้อมที่กำหนด โดยแกนนอนคือขนาดของสัญญาณที่ได้รับ และแกนตั้งคือค่าความน่าจะเป็นที่สามารถรับกำลังสัญญาณในระดับนั้นได้ ซึ่งมีรายละเอียดดังนี้



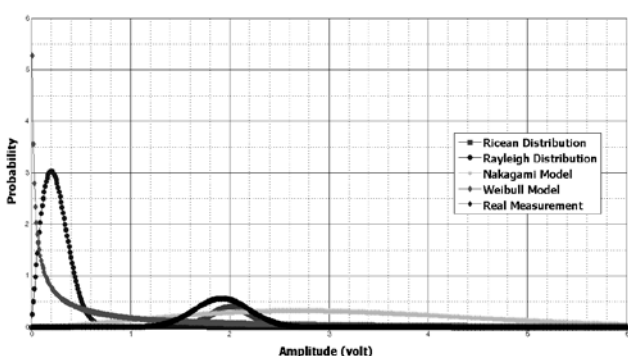
ภาพที่ 5: แสดงผลการทดลองการจางหายของกำลังสัญญาณในสภาวะแวดล้อมแบบพื้นที่เปิดโล่ง



ภาพที่ 6: แสดงผลการทดลองการจางหายของกำลังสัญญาณในสภาวะแวดล้อมแบบชุมชนเมือง



ภาพที่ 7: แสดงผลการทดลองการจางหายของกำลังสัญญาณในสภาวะแวดล้อมแบบชานเมือง



ภาพที่ 8: แสดงผลการทดลองการจางหายของกำลังสัญญาณในสภาวะแวดล้อมแบบชานเมือง

5. สรุปผลการวิจัยและข้อเสนอแนะ

สืบเนื่องจากผลการทดลองที่ได้กล่าวในหัวข้อ 4.1 และ 4.2 จะเห็นได้ว่าแนวโน้มการเกิดการลดทอนตามระยะทางจะเพิ่มมากขึ้นตามระยะห่างระหว่างเครื่องส่งและเครื่องรับที่เพิ่มมากขึ้น

ขึ้น ส่วนแนวโน้มการเกิดการจางหายของกำลังสัญญาณจะมีความแตกต่างกันตามแต่ละแบบจำลองและมีโอกาสที่จะได้รับกำลังสัญญาณสูงสุดในปริมาณที่แตกต่างกันอีกด้วย ซึ่งจากการวิเคราะห์และเปรียบเทียบผลการทดลองจากการเลียนแบบโดยโปรแกรมคอมพิวเตอร์กับผลการทดลองในสภาวะแวดล้อมจริงทำให้สามารถสรุปแบบจำลองที่มีความเหมาะสมที่สุดที่จะนำไปใช้คาดการณ์การลดทอนตามระยะทางและการจางหายของกำลังสัญญาณในแต่ละสภาวะแวดล้อมภายในประเทศไทยได้ดังตารางที่ 2 และตารางที่ 3 ตามลำดับ

ตารางที่ 2: แสดงแบบจำลองที่มีความเหมาะสมสำหรับการนำไปใช้พยากรณ์การลดทอนตามระยะทางในสภาวะแวดล้อมต่างๆ

สภาพแวดล้อม	แบบจำลองที่เหมาะสม
พื้นที่เปิดโล่ง	Lee Model
ชุมชนเมือง	Ericsson Model
ชานเมือง	Erceg Model
ชนบท	COST-231 Model

ตารางที่ 3: แสดงแบบจำลองที่มีความเหมาะสมสำหรับการนำไปใช้พยากรณ์การจางหายของกำลังสัญญาณในสภาวะแวดล้อมต่างๆ

สภาพแวดล้อม	แบบจำลองที่เหมาะสม
พื้นที่เปิดโล่ง	Ricean Distribution
ชุมชนเมือง	Rayleigh Distribution
ชานเมือง	Nakagami Model
ชนบท	Ricean Distribution

เอกสารอ้างอิง

- [1] Winton Afric, Branka Zovko-Cihlar and Sonja Grgic, "Methodology of Path Loss Calculation using Measurement Result," *6th EURASIP Conference focused on Speech Image Processing and Multimedia Communication and Services*, Maribor, Slovenia, June 27-June 30, pp.257-260, 2007
- [2] Mohammed Alsami, Tughrul Arslan, John Thompson and Ahmet T. Erdogan, "Frequency Analysis of Path Loss Models on WiMAX," *3rd Computer Science and Electronic Engineering Conference (CEEC)*, Colchester, United Kingdom, 2011.
- [3] Liu Yarong, Yang Xiaofei, Shen Zhoujun and Lu Qiu, "The Research of Path Loss Model Based on WiMAX Wireless Channel," *2nd International Conference on*

Communication Systems, Damai Beach Resort, Kuching, Malaysia, 2010.

- [4] Hammant Kumar Sharma, Sanjeev Sharma and Krishna Pandey, "Survey of propagation Model in Wireless Network," *IJCSI International Journal of Computer Science Issues*, Vol.8, Issue 3, No. 2, pp. 468-472, 2011.
- [5] Tapan K. Sakar, Zhong Ji, Kyungjung Kim, Abdellatif Medouri and Magdalena Salazar-Palma, "A Survey of Various Propagation for Mobile Communication," *IEEE Antennas and Propagation Magazine*, Vol. 45, No. 3, pp. 51-82, 2003.

การตรวจจับการปลอมแปลงโพรโทคอลเออาร์พี โดยค่าแฮชของ MAC Address Gateway ARP Poisoning Detection by MAC Address Gateway's Hash value

หัตสดา พฤฒาจารย์บอดิ (Hatsada Phruekthachanbodi)¹ และศิริปัฐ บุญครอง (Sirapat Boonkrong)²

^{1,2}ภาควิชาการสื่อสารข้อมูลและเครือข่าย คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹palm.dn55@gmail.com, ²sirapatb@kmutnb.ac.th

บทคัดย่อ

การปลอมแปลงโพรโทคอลเออาร์พี (ARP Poisoning) เป็นปัญหาสำคัญที่อาจจะก่อให้เกิดความเสียหายต่อผู้ใช้งานระบบอินเทอร์เน็ตได้ การโจมตีหลักๆมีอยู่ด้วยกัน 2 แบบ คือ DOS (Denial of service) จะทำให้เครื่องที่ถูกโจมตีไม่สามารถใช้งานเครือข่ายได้ และการโจมตีอีกวิธีหนึ่งคือ MITM (Man in the middle) ผู้โจมตีจะแทรกเข้าไปอยู่ตรงกลางระหว่างเครื่องที่ถูกโจมตีและเครือข่าย งานวิจัยนี้มีวัตถุประสงค์ประสงค์เพื่อทำการตรวจจับการปลอมแปลงโพรโทคอลเออาร์พี โดยใช้ค่า Hash ของ MAC Address Gateway เพื่อเป็นการเปรียบเทียบความถูกต้องของ MAC Address Gateway ที่ใช้งานจริง โดย Application ที่ผู้วิจัยทำการพัฒนาขึ้นมา ซึ่งจะใช้งานเครือข่ายแบบ DHCP ดังนั้น Application ที่ผู้วิจัยพัฒนาขึ้นมาที่จะทำการตรวจจับการโจมตีแบบ rogue DHCP ได้อีกด้วย ซึ่ง application สามารถใช้งานได้ทั้ง ระบบปฏิบัติการ windows และ linux โดยไม่ต้องปรับแก้โพรโทคอล ARP และ DHCP

คำสำคัญ: เออาร์พี แฮชแมค อาร์พพ้อยโซนิง ดีแฮชซีพี ไรค์ดีแฮชซีพี

Abstract

ARP poisoning is an attack that can cause damages to users' application and network. There are two main categories of attack : Denial of Service (DOS), where the victim is not able to use or access the network , and Man-in-the-Middle (MITM) where an attacker cause in between the network and the victim. The objective of this research is to develop a method and an application that can detect ARP poisoning by comparing the hash values of the MAC address gateway. The developed application

is to be used on a DHCP network. Not any that the application can detect ARP poisoning , it can detect rogue DHCP as well.

Keyword: ARP, HMAC, ARP Poisoning, DHCP, rouge DHCP

1. บทนำ

การติดต่อสื่อสารกันในระบบเครือข่าย LAN (Local Area Network) จะใช้ ARP (Address Resolution Protocol) [1] ในการค้นหา MAC Address ซึ่งจะทำการจับคู่ระหว่าง IP Address และ MAC Address เพื่อใช้ในการติดต่อสื่อสารกัน แต่การทำงานของ [2] ARP (Address Resolution Protocol) ยังคงมีความเสี่ยงที่จะถูกโจมตีแบบ ARP Poisoning ได้เพราะปัญหาการ Update ARP Cache ใน ARP (Address Resolution Protocol) จะไม่มีการตรวจสอบว่าค่า MAC Address ที่ได้รับมามีความถูกต้องหรือไม่ ซึ่งในการโจมตีแบบ ARP Poisoning นั้นสามารถทำได้ง่าย โดยผู้ที่ทำการโจมตีนั้นไม่จำเป็นต้องมีความรู้ทางด้านคอมพิวเตอร์มากนัก พร้อมทั้งมีเครื่องมือหรือโปรแกรมที่ใช้ในการโจมตีดังกล่าวจำนวนมากเช่น [3] NetCut Switch Sniffer Cain & Abel Backtrack Ettercap NG และอื่น ๆ จึงเป็นการง่ายสำหรับ ผู้โจมตีมือสมัครเล่น ที่จะทำการโจมตีและสร้างความเสียหายให้กับระบบเครือข่าย

ปัจจุบันได้มีซอฟต์แวร์และงานวิจัยที่จะแก้ไขปัญหา การปลอมแปลง โพรโทคอลเออาร์พีมาก่อนแล้วจำนวนหนึ่ง เช่น Secure ARP (S-ARP) Anti NetCut Dynamic ARP Inspection Anti ARP เป็นต้น หรือการใช้ฮาร์ดแวร์ ในการป้องกัน [4] แต่จากการศึกษาพบว่าการแก้ปัญหาเหล่านี้ล้วนมีความไม่สมบูรณ์ มีค่าใช้จ่ายสูง หรือไม่เหมาะสมในการนำไปใช้จริง งานวิจัยนี้

ได้ออกแบบการตรวจจับการปลอมแปลงโพรโทคอลเออาร์พี โดยการ Hash แบบ HMAC ซึ่งจะทำการ Hash MAC Address Gateway โดยใช้ความสามารถของ DHCP Server คือ DHCP Option 15 ในการส่งค่า Hash Value ไปยังเครื่อง Client ซึ่งผู้วิจัยได้ทำการพัฒนา Application ที่จะติดตั้งในฝั่งของ Client เพื่อทำการรับค่า Hash Value จาก DHCP Server option 15 มาทำการตรวจสอบความถูกต้องและแจ้งแจ้งเตือนไปยังผู้ใช้งานว่าเครือข่ายที่ใช้งานอยู่มีความปลอดภัยหรือไม่

บทความนี้ประกอบไปด้วยส่วนประกอบต่างๆดังนี้ หัวข้อที่ 2 จะกล่าวถึง ทฤษฎีและงานวิจัยที่เกี่ยวข้อง หัวข้อที่ 3 จะกล่าวถึงวิธีดำเนินการวิจัย หัวข้อที่ 4 จะกล่าวถึงผลการดำเนินงาน และหัวข้อที่ 5 จะกล่าวถึงบทสรุปของงานวิจัย

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 โพรโทคอลเออาร์พี ARP

ARP (Address Resolution Protocol) [5] เป็นโพรโทคอลที่ใช้ในการค้นหา MAC Address และจับคู่ระหว่าง MAC Address กับ IP Address เพื่อใช้ในการติดต่อสื่อสารกันบนเครือข่าย การจะสื่อสารกันเครื่องต้นทางจะทราบเพียง IP Address ของเครื่องปลายทางซึ่งจะไม่สามารถสื่อสารกันได้เลยจึงต้องรู้ MAC Address ของเครื่องปลายทางด้วย โพรโทคอล ARP จึงทำหน้าที่ในการค้นหา MAC Address ของเครื่องปลายทาง โดยจะทำการส่ง ARP Request แบบ broadcast ไปยังทุกเครื่องใน Network เดียวกัน โดยจะทำการส่ง IP Address และ MAC Address ของเครื่องต้นทางและ IP Address ของเครื่องปลายทางไปด้วยเมื่อเจอเครื่องปลายทางก็จะทำการนำ MAC Address ของเครื่องต้นทางมา Update ARP Cache โดยไม่มีการตรวจสอบ และทำการส่งค่า MAC Address ไปยังเครื่องต้นทาง เครื่องต้นทางก็จะทำการ Update ARP Cache โดยไม่มีการตรวจสอบเช่นกัน จุดอ่อนของโพรโทคอล ARP ก็คือการ Update ARP Cache โดยไม่มีการตรวจสอบว่า MAC Address นั้นถูกต้องหรือไม่ ทำให้เป็นช่องโหว่ที่สามารถถูกโจมตีได้ง่าย

2.2 การโจมตีแบบ ARP Poisoning

ปัจจุบันการโจมตีแบบ ARP Poisoning [6] ส่งผลกระทบต่อระบบเครือข่าย LAN ซึ่งผู้โจมตี จะโจมตีจุดอ่อนของโพรโทคอล ARP คือเมื่อเครื่องคอมพิวเตอร์มีการรับ ARP

Request หรือ ARP Reply ก็จะมีการ Update ARP Cache ทันที [7] ไม่มีการตรวจสอบว่าข้อมูลนั้นมาจากไหน เป็นจริงหรือไม่

หากผู้โจมตีทำการส่ง ARP Request ไปยังเครื่องเหยื่อโดยส่ง IP Address ของ Gateway และ Mac ที่ไม่ถูกต้องไปยังเครื่องเหยื่อส่งผลกระทบต่อเครื่องเหยื่อไม่สามารถติดต่อกับ Gateway ได้ ทำให้เครื่องเหยื่อไม่สามารถออกใช้งานอินเทอร์เน็ตได้ซึ่งเรียก การโจมตีแบบนี้ว่า เป็นแบบ Denial of service [1]

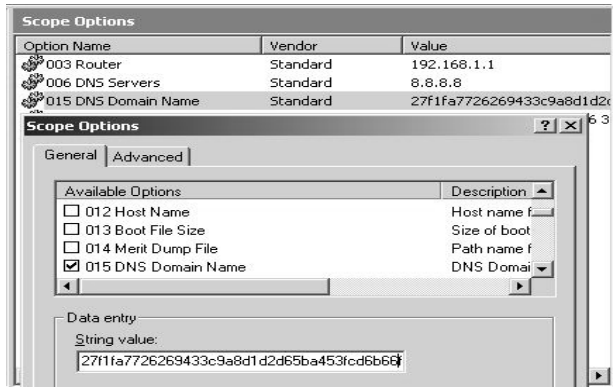
หรือ ผู้โจมตีทำการส่ง ARP Reply ไปยังเครื่องเหยื่อเพื่อบอกว่า เครื่องผู้โจมตีเป็น Gateway และส่ง ARP Reply ไปยัง Gateway เพื่อบอก Gateway ว่า เครื่องผู้โจมตีเป็น เครื่องเหยื่อ ทำให้เวลาที่มีการติดต่อสื่อสารกันข้อมูลทุกอย่างจะผ่านเครื่องผู้โจมตี ส่งผลให้ข้อมูลทุกอย่างอาจจะถูกดักจับได้ ซึ่งเรียกการโจมตีแบบนี้ว่า เป็นแบบ Man in the Middle [1]

2.3 HMAC

HMAC คือ Hash Function ประเภทหนึ่ง [8], [9] ซึ่งการทำงานคือ การนำเอา plaintext เช่นไฟล์เอกสาร หรือไฟล์อื่นๆ มาผ่านกระบวนการทางคณิตศาสตร์ที่เรียกว่า Hash function เพื่อทำการย่อข้อมูลให้อยู่ในรูปของ ciphertext หรือ hash value การทำงานของ Hash function จะมีความเร็วสูง และ HMAC ต้องใช้ key ในการในการย่อข้อมูล เพื่อให้ค่า hash value มีความปลอดภัยมากขึ้น โดยในงานวิจัยนี้จะนำค่า MAC address gateway มาทำการ HMAC และจะนำค่า hash value ที่ได้ไปเปรียบเทียบกับผลลัพธ์ตาม algorithm ของ application

2.4 DHCP

DHCP (Dynamic Host Configuration Protocol) คือ Protocol ที่กำหนด IP Address [10], [11] แบบอัตโนมัติแก่เครื่องลูกข่าย ในใช้งานระบบเครือข่ายแบบ DHCP Server โดย DHCP Server จะทำการแจกจ่าย IP Address, Subnet mask, Gateway, IP Address DNS, กำหนดอายุการใช้งานของ IP Address ที่ใช้ อยู่ และ DHCP option ที่มีอยู่ทั้งหมด 255 option แต่ไม่ได้เปิดใช้งานทั้งหมด ซึ่งในงานวิจัยนี้ได้เลือก DHCP option 15 มาทำการส่งค่า hash value ไปยัง client โดยปกติของ option 15 [12] จะไม่ได้เปิดมาใช้ งาน แต่การใช้งานส่วนใหญ่จะเป็นการกำหนดค่า Parent domain name เพื่อบอกว่าองค์กรใช้ domain name ชื่ออะไร การกำหนดค่า hash value จะแสดงตามภาพที่ 1



ภาพที่ 1: การใส่ค่า hash value ใน option 15

2.5 การทบทวนวรรณกรรม

การวิเคราะห์ศึกษาและเปรียบเทียบสมรรถภาพของวิธีการป้องกัน ARP Spoof [13] ซึ่งมีการนำเสนอวิธีที่น่าสนใจคือการ Test base โดยการนำโปรแกรมป้องกันการโจมตีแบบ ARP Spoof ที่ใช้กันอยู่หลากหลายมาเปรียบเทียบโดยการสร้างสถานการณ์จริงขึ้นมา ซึ่งผลการทดสอบ คือ โปรแกรมสามารถป้องกันได้แค่ อย่างใดอย่างหนึ่ง คือในหนึ่งโปรแกรม ไม่สามารถป้องกันได้ทั้ง DOS และ MITM ได้ [14] ระบบป้องกันและเตือนภัยการปลอมแปลงโปรโตคอลเออาร์พีประเภทซอฟต์แวร์ โปรแกรมเพื่อการประยุกต์ใช้งานสำหรับลินุกซ์ โดยมีรายละเอียดดังนี้ Vaccine Server (VS) เป็นโปรแกรมที่ติดตั้งบนเครื่อง Server ทำหน้าที่เก็บ Static ARP entry ระหว่างเครื่องที่เป็น Vaccine Gateway (VG) กับ Client และ (VS) ยังมีหน้าที่ในการเตือนภัยไปยัง Vaccine Client (VC) Vaccine Client จะเป็นโปรแกรมที่ติดตั้งสำหรับ Client ทุกเครื่อง แต่การใช้ระบบ จะมีความยุ่งยากเพราะ ยังต้องมี Server เพื่อติดตั้งโปรแกรมการแจ้งเตือน และต้องใช้ Gateway ที่เป็น Server เพราะจะต้องติดตั้ง โปรแกรมลงบน Gateway ทำให้ระบบจะยังมีความซับซ้อน [3] ระบบป้องกันการปลอมแปลงโปรโตคอล ARP แบบพลวัตในเครือข่ายเอสเอ็มอีจริง ซึ่งใช้ DHCP Server ในการแจกจ่าย IP address เมื่อ client ได้รับ MAC address gateway ก็จะมีการสร้าง static ARP เพื่อป้องกัน แต่หากถูกโจมตีแบบ rogue DHCP ระบบจะไม่สามารถป้องกันได้ [15] Secure ARP ออกแบบโปรโตคอลเออาร์พีขึ้นมาใหม่ โดยอาศัยกลไกการเข้ารหัสและลงลายเซ็นดิจิทัล (Digital signature) เพื่อยืนยันความถูกต้องของโปรโตคอล ซึ่ง S-ARP ถูกแก้ไขในระบบปฏิบัติการ ลินุกซ์ แล้วสามารถป้องกันได้ อย่างไรก็ตามวิธีนี้ไม่สามารถนำไปใช้ได้กับ

โปรโตคอลเออาร์พีเดิมที่มีอยู่ เพราะเป็นการแก้ไขโปรโตคอลเออาร์พีใหม่ ซึ่งอาจเป็นไปได้ [5] เป็นการสร้างเครือข่ายเฉพาะสำหรับการติดต่อสื่อสารกันภายใน โดยการออกแบบโปรโตคอลขึ้นมาใหม่เพื่อหลีกเลี่ยงการถูกโจมตีแบบ ARP poisoning จึงทำให้มีข้อจำกัดในการใช้งานเพราะจะใช้งานได้เฉพาะเครือข่ายที่ทำงานในโปรโตคอลนี้เท่านั้น

3. วิธีดำเนินการวิจัย

3.1 การเลือก Cryptographic hash function

ในงานวิจัยนี้ได้ทำการทดลองการ Hash MAC address gateway ด้วย Cryptographic hash function แบบต่างๆดังนี้ MD5, SHA1, HMAC-MD5, เพื่อที่จะเลือก Cryptographic ที่เหมาะสมที่สุด ซึ่งในงานวิจัยนี้จะเลือก HMAC-SHA1 เพราะมี Output 160bit ซึ่งจะมีโอกาสเกิด collision น้อยกว่า MD5 ที่ 128bit HMAC มี Keyในการทำงานเพื่อเพิ่มความซับซ้อนให้กับ Output และระยะเวลาในการทำงานไม่ต่างจาก Hash function ตัวอื่น ทดสอบ Hash MAC address gateway จำนวน 500 รอบ ตารางที่ 1 จะแสดง คุณสมบัติของ Hash Function แต่ละแบบ

ตารางที่ 1: เปรียบเทียบ คุณสมบัติของ Hash Function แต่ละแบบ

	MD5	SHA1	HMAC-MD5	HMAC-SHA1
Out put	128 bit	160 bit	128 bit	160 bit
Use key	No	No	Yes	Yes
Time	0.010584 (ms)	0.010584(ms)	0.010584(ms)	0.010584(ms)

3.2 การออกแบบและพัฒนาระบบ

3.2.1 แนวคิดการออกแบบระบบ

ภาพรวมของระบบในงานวิจัยนี้เป็น การนำเสนอการตรวจจับการปลอมแปลงโปรโตคอลเออาร์พีโดยการเปรียบเทียบค่า Hash (MAC Address Gateway) ที่ได้รับมาจาก DHCP Server ซึ่งจะเป็นค่า Hash ที่ถูกต้อง และค่า Hash (MAC Address Gateway) ของเครื่อง Client ที่อาจจะถูกปลอมแปลงได้ เพื่อเปรียบเทียบว่าค่า Hash ที่ Client ใช้อยู่มีค่าตรงกับค่า Hash ที่ได้รับจาก DHCP Server หรือไม่

3.2.2 รายละเอียดการพัฒนาระบบ

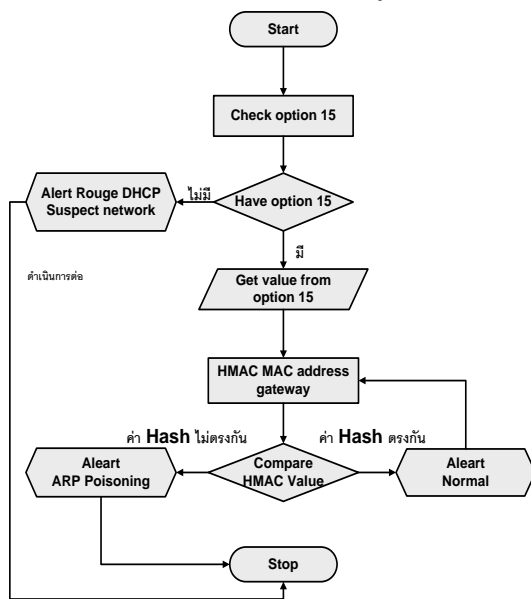
ระบบตรวจจับการปลอมแปลงโปรโตคอลเออาร์พีโดยการเปรียบเทียบค่า Hash ซึ่งจะมีรายละเอียดการทำงานของระบบคือ ทำการ HMAC ค่า MAC Address Gateway ที่ใช้งานอยู่จริง โดยผู้วิจัยจะทำการใช้โปรแกรม Open SSL ในการสร้าง

ค่า Hash และเมื่อได้ค่า Hash จะทำการเปิด DHCP Option 15 เพื่อบรรจุค่า Hash ลงไป ซึ่งจะไว้ สำหรับส่งไปให้ Client ทำการตรวจสอบ ซึ่งในส่วนของ DHCP Server จะมีขั้นตอนเพียงเท่านี้

ในการทำงานของ Algorithm ใน Application ที่จะใช้ตรวจสอบเพื่อตรวจจับการโจมตี โดย Application จะเริ่มจากการตรวจเช็ค ค่า Hash ว่าได้รับแล้ว หรือไม่มีการส่งค่า Hash มา ในขั้นตอนนี้หากไม่ได้รับค่า Hash มา ระบบจะทำการแจ้งเตือนว่า เป็นการโจมตีแบบ rogue DHCP เพราะ DHCP Server ที่ผู้วิจัยนำมาใช้งานจะต้องมีค่า Hash ที่ได้นำมาบรรจุไว้ แต่ถ้าหากมีค่า Hash จะทำการเก็บค่า Hash ไว้ ขั้นตอนต่อไป

ระบบจะทำการ HMAC ค่า MAC Address Gateway ที่ใช้อยู่ เมื่อได้ค่า Hash ออกมา ก็จะนำไปเปรียบเทียบกับค่า Hash ที่ได้รับจาก DHCP Server Option 15 ว่ามีค่าตรงกันหรือไม่ หากค่า Hash ไม่ตรงกัน จะแจ้งเตือนผู้ใช้งานว่า มีการโจมตีแบบ ARP Poisoning แต่ถ้าค่า Hash ตรงกันระบบจะแจ้งเตือนผู้ใช้งานว่าเครือข่ายที่ใช้อยู่มีความปลอดภัย และโปรแกรมจะทำงานต่อไปในขั้นตอนของการ Hash ค่า MAC Address Gateway ที่ใช้อยู่และนำมาเปรียบเทียบกับค่า Hash ที่ได้รับมาจาก DHCP Server Option 15 จนกว่าจะถูกโจมตี และจะทำการแจ้งเตือนไปยังผู้ใช้งาน

รายละเอียดขั้นตอนการทำงานสามารถดูได้จากภาพที่ 2



ภาพที่ 2: Algorithm for application

4. ผลการดำเนินงาน

งานวิจัยนี้ได้ออกแบบ Algorithm การตรวจจับการปลอมแปลง โพรโทคอลเออาร์พี ซึ่งใช้แอปพลิเคชันการเปรียบเทียบค่า Hash ของ MAC Address Gateway และใช้เทคนิคการส่งค่า Hash ที่ถูกต้องด้วย DHCP Option 15 เพื่อนำค่า Hash ไปเปรียบเทียบ โดย Application ที่ได้พัฒนาขึ้นมาเพื่อทำงานตาม Algorithm ที่ผู้วิจัยทำการออกแบบและพัฒนา

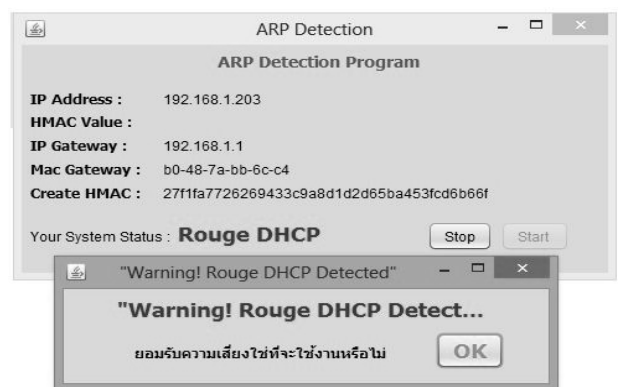
4.1 ทดสอบการโจมตี

การทดสอบการโจมตีนั้น ในงานวิจัยนี้จะแบ่งการโจมตีออกเป็น 2 แบบ แบบที่ 1 rogue DHCP เป็นการปลอม DHCP Server เพื่อแจก IP address ที่ไม่ถูกต้องส่งผลกระทบต่อให้เกิดการโจมตีแบบ DOS MITM และ DNS Spoofing [17] แบบที่ 2 ARP Poisoning คือ จะทำให้เครื่องที่ถูกโจมตีไม่สามารถใช้งาน เครือข่ายได้เลย โดยจะใช้โปรแกรม Net cut ทำการโจมตี หรือ ผู้โจมตีจะเข้าแทรกไปอยู่ตรงกลางระหว่างเครื่องเหยื่อกับเครือข่าย เพื่อดักจับข้อมูล โดยจะใช้โปรแกรม cain & abel เพื่อทำการโจมตี

4.2 การตรวจจับการโจมตี

4.2.1 ตรวจจับการโจมตี rogue DHCP

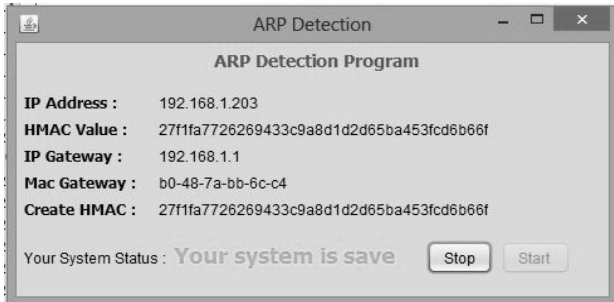
ในการโจมตีแบบ rogue DHCP คือ จะใช้ DHCP Server ที่ไม่ได้เปิด DHCP Option 15 มาทำการแจก IP address ให้เครื่อง client ผลของการตรวจจับคือ Application จะทำการรับค่า Hash จาก DHCP Option 15 เมื่อไม่มีค่า Hash ส่งมายัง Application ก็จะมีการแจ้งเตือนผู้ใช้งานทันที เพราะ DHCP Server ที่ผู้วิจัยนำมาใช้ในระบบจะต้องมีค่า Hash จาก Option 15 แบบมาด้วย หน้าตาของ application เมื่อตรวจพบการโจมตีจะแสดงตามภาพที่ 3



ภาพที่ 3: แสดงการ Detected rogue DHCP

4.2.2 การทำงานเมื่อไม่มีการโจมตี

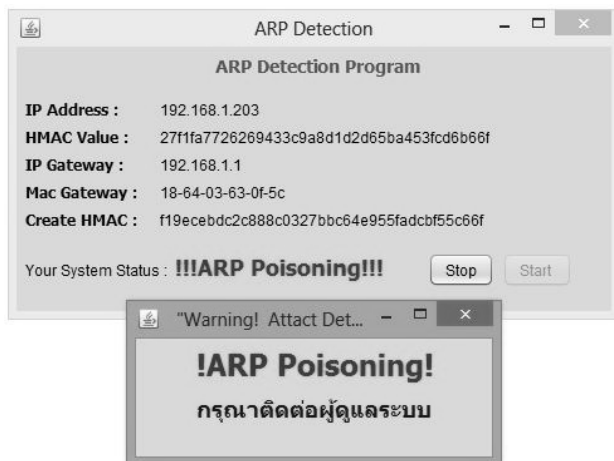
สำหรับการทำงานของ Application เมื่อไม่มีการโจมตี ก็จะได้รับค่า Hash จาก DHCP Option 15 และทำการ Hash ค่า MAC Address Gateway ที่ใช้อยู่ ถ้ามีค่า hash value เท่ากัน Application จะทำการบอกผู้ใช้งานว่า “Your system is save” และ Application จะทำ Algorithm การตรวจจับการโจมตีอยู่ตลอดเวลาจนกว่าจะพบการโจมตี หน้าตาโปรแกรมเมื่อไม่มีการโจมตีจะแสดงดังภาพที่ 4



ภาพที่ 4: แสดงการทำงานเมื่อไม่มีการโจมตี

4.2.3 การตรวจจับ ARP Poisoning

การโจมตีแบบ ARP poisoning ผู้โจมตีจะใช้โปรแกรม netcut ในการโจมตีแบบ DOS และใช้โปรแกรม cain&abel ในการโจมตีแบบ MITM [13] ซึ่งทั้ง 2 วิธีจะเป็นการปลอมแปลงค่า MAC address gateway การทำงานของ Application เมื่อตรวจพบความผิดปกติของ MAC Address Gateway ที่ใช้อยู่ และค่า Hash ของ MAC address gateway มีความแตกต่างกับค่า Hash ที่ได้รับมาจาก DHCP Option 15 แสดงว่ามีการโจมตีแบบ ARP Poisoning เกิดขึ้น Application จะทำการแจ้งเตือนผู้ใช้งานทันที โดยแสดงดังภาพที่ 5



ภาพที่ 5: แสดงการ Detected ARP Poisoning

4.3 การวิเคราะห์ประสิทธิภาพของโปรแกรม

ผลทดลองการโจมตีในงานวิจัยนี้ได้ทดลองการโจมตีแบบ ARP Poisoning ในรูปแบบของ Denial of service และ Man in the middle จำนวน 30 ครั้ง และ ทาการทดลองการโจมตีแบบ rogue DHCP จำนวน 30 ครั้ง application สามารถตรวจจับและแจ้งเตือนการโจมตีได้ทุกครั้ง แสดงในตารางที่ 2 และทำการเปิด application ไว้ในสถานการณ์ปกติ ที่ไม่มีการโจมตีโดยใช้เวลาในการทดสอบ 1 ชั่วโมง application สามารถทำงานได้ปกติโดยไม่มีการแจ้งเตือนที่ผิดพลาด

ตารางที่ 2: แสดงการตรวจจับการโจมตี

Attack	Detected
Rogue DHCP	100%
ARP Poisoning	100%

จะเห็นได้ว่า Algorithm การตรวจจับสามารถ ตรวจจับการโจมตีได้ทั้งแบบ rogue DHCP DOS และแบบ MITM โดยการเปรียบเทียบค่า hash value ซึ่งในงานวิจัย [13] จะพบว่า โปรแกรมป้องกันสามารถกันได้แค่ DOS หรือ MITM เท่านั้น ส่วนงานวิจัย [14] จะมีความยุ่งยากในการใช้งานเพราะการสื่อสารกันจะมีการเข้ารหัสซึ่งทำให้ระบบมีความล่าช้าและยังต้องเพิ่ม Server ขึ้นมาเพื่อแจ้งเตือน งานวิจัย [3] สามารถตรวจจับได้ทั้ง 2 แบบเหมือนกัน แต่ระบบ ยังไม่มีการตรวจจับ rogue DHCP ในงานวิจัย [15] จะต้องมีการแก้ไขเปลี่ยนแปลง โพรโทคอลเออาร์พี ซึ่งจะเพิ่มความยุ่งยากในการนำไปใช้งานจริง และงานวิจัย [5] จะเป็นการออกแบบโพรโทคอลในการสื่อสารขึ้นมาใหม่ ซึ่งไม่สามารถนำไปใช้ในเครือข่าย LAN ทั่วไปได้

5. สรุป

งานวิจัยนี้ได้เสนอ Algorithm การตรวจจับการปลอมแปลง โพรโทคอลเออาร์พีโดยการเปรียบเทียบค่า Hash ของ MAC Address Gateway ผลการทดลองได้พบว่า Application มีประสิทธิภาพในการตรวจจับการโจมตีได้ทั้งแบบ DOS และ MITM และในเครือข่ายแบบ DHCP Application ยังสามารถตรวจจับ rogue DHCP ได้อีกด้วย งานวิจัยนี้สามารถตรวจจับ การขัดขวางการเชื่อมต่อไปยังเครือข่าย และ การถูกดักจับข้อมูล

ซึ่งการโจมตีแบบดักจับข้อมูล จะทำให้ข้อมูลรั่วไหลไปยังผู้ไม่หวังดีได้ นำไปใช้ได้ทั้งในระบบปฏิบัติการ Windows และ Linux สามารถนำไปใช้ได้จริงโดยไม่ต้องแก้ไขเปลี่ยนแปลง โพรโทคอลเออาร์พี ข้อจำกัดของงานวิจัยคือ ต้องใช้เครือข่ายแบบ DHCP เท่านั้นเพื่อทำการส่งค่า Hash ไปยัง Client ในอนาคตสามารถพัฒนาให้ระบบสามารถสร้าง Static ARP หลังการตรวจสอบความถูกต้องเพื่อเป็นการป้องกันการโจมตีได้

เอกสารอ้างอิง

- [1] G. Jinhua and X. Kejian, "ARP spoofing detection algorithm using ICMP protocol," in *2013 International Conference on Computer Communication and Informatics (ICCCI)*, 2013, pp. 1–6.
- [2] S. Puangpronpitag and N. Masusai, "An efficient and feasible solution to ARP Spoof problem," in *6th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*, 2009. ECTI-CON 2009, 2009, vol. 02, pp. 910–913.
- [3] ณรงค์ฤทธิ์ มะสุใส, สมนึก พ่วงพรพิทักษ์. ระบบป้องกันการปลอมแปลงโพรโทคอล ARP แบบพลวัตในเครือข่ายเอสเอ็มอีจริง, 2553.
- [4] N. Agrawal, B. Pradeepkumar, and S. Tapaswi, "Preventing ARP spoofing in WLAN using SHA-512," in *2013 IEEE International Conference on Computational Intelligence and Computing Research (ICCCIC)*, 2013, pp. 1–5.
- [5] W. Xiao-ling and Z. Gang, "Defence Design for ARP Spoofing Based on NDIS Intermediate Driver," in *2012 International Conference on Computer Distributed Control and Intelligent Environmental Monitoring (CDCIEM)*, 2012, pp. 218–220.
- [6] S. G. Bhirud and V. Katkar, "Light weight approach for IP-ARP spoofing detection and prevention," in *2011 Second Asian Himalayas International Conference on Internet (AH-ICI)*, 2011, pp. 1–5.
- [7] H. Salim, Z. Li, H. Tu, and Z. Guo, "Preventing ARP Spoofing Attacks through Gratuitous Decision Packet," in *2012 11th International Symposium on Distributed Computing and Applications to Business, Engineering Science (DCABES)*, 2012, pp. 295–300.
- [8] S. W. Jung and S. Jung, "HRP: A HMAC-based RFID mutual authentication protocol using PUF," in *2013 International Conference on Information Networking (ICOIN)*, 2013, pp. 578–582.
- [9] F. Zhang and Z. J. Shi, "Differential and Correlation Power Analysis Attacks on HMAC-Whirlpool," in *2011 Eighth International Conference on Information Technology: New Generations (ITNG)*, 2011, pp. 359–365.
- [10] C. Lin, T. Su, and Z. Wang, "Summary of high-availability DHCP service solutions," in *2011 4th IEEE International Conference on Broadband Network and Multimedia Technology (IC-BNMT)*, 2011, pp. 12–17.
- [11] ปิยวัฒน์ คำสบาย, สมนึก พ่วงพรพิทักษ์. เจอาร์พี: ซอฟต์แวร์ป้องกันการปลอมแปลงโพรโทคอล ARP. มหาวิทยาลัยมหาสารคาม,
- [12] คู่มือ windows server 2003 R2 โปรวิชั่น กรุงเทพฯ
- [13] นายภวิษ ขุนใหญ่. การวิเคราะห์ศึกษาและเปรียบเทียบสมรรถภาพของวิธีการป้องกัน ARP SPOOF. มหาวิทยาลัยขอนแก่น, 2553.
- [14] ธงชัย เจือจันทร์, ปิยวัฒน์ คำสบาย. ระบบป้องกันและเตือนภัยการปลอมแปลงโพรโทคอลเออาร์พีประเภทซอฟต์แวร์ โปรแกรมเพื่อการประยุกต์ใช้งานสำหรับลินุกซ์, 2551.
- [15] D. Bruschi, A. Ornaghi, and E. Rosti, "S-ARP: a Secure Address Resolution Protocol," University degli Studi di Milano, Italy, 2012.
- [16] C. Lai, K.-C. Leung, and V. O. K. Li, "Design and analysis of TCP AIMD in wireless networks," in *2013 IEEE Wireless Communications and Networking Conference (WCNC)*, 2013, pp. 1422–1427.
- [17] A. Amin, H. Ahmed, A. Rafique, M. Junaid Nawaz, M. Salahudin, and Z. Ahmed, "Designing a hierarchical DHCP server's model to automatically provide dedicated IP address anywhere in the world with mobility," in *Internet Technology and Secured Transactions (ICITST)*, 2011 International Conference for, 2011, pp. 250–253.

ผลกระทบของการรบกวนเครือข่ายไร้สายแบบพัลส์ และแบบกวาด ที่มีต่อเครือข่ายไร้สายแบบ Ad hoc

The Effective of Wireless Jamming: Pulsed Jamming and Swept Jamming for Wireless Ad hoc Network

เอกรัฐ หล่อพิเชียร (Aekkarat Lorphichian)¹ และ ศิริปัฐ บุญครอง (Sirapat Boonkrong)²

¹สาขาวิชาเทคโนโลยีสารสนเทศการศึกษาคณะครุศาสตร์อุตสาหกรรม มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

²คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹aekkarat@mail.rmutt.ac.th, ²sirapatb@kmutnb.ac.th

บทคัดย่อ

การรบกวนเครือข่ายไร้สาย (Wireless Jamming) เป็นภัยคุกคามรูปแบบหนึ่งของ DoS (Denial of Service) ซึ่งผลการรบกวนทำให้ประสิทธิภาพการทำงานของเครือข่ายไร้สายลดลง หรืออาจจะทำให้เครือข่ายไร้สายไม่สามารถทำงานได้ งานวิจัยนี้ได้นำเสนอผลกระทบของการรบกวนเครือข่ายไร้สายแบบพัลส์และแบบกวาดที่มีต่อค่าเฉลี่ยทราฟฟิค (Average Throughput) ของเครื่องที่ถูกรบกวน โดยตัวแปรที่ศึกษาคือ กำลังตัวส่งของผู้รบกวน (Jammer Transmission Power) เวลาที่รบกวน และ จำนวนผู้รบกวน

คำสำคัญ: การรบกวนเครือข่ายไร้สาย, การรบกวนแบบพัลส์, การรบกวนแบบกวาด

Abstract

Wireless Jamming is a type of Denial of Service (Dos) attacks that can reduce the performance of the network and even stop it from working in some cases. This research provides a study of impact from two types of wireless jamming on a wireless ad hoc network. The jamming attacks studied here include pulsed jamming and swept jamming. The average throughput is studied based on the changes of different parameters: jammer transmission power, duration of jamming and the number of jammers.

Keyword: wireless jamming, pulsed jamming, swept jamming.

1. บทนำ

เครือข่ายไร้สายแบบ Ad hoc (Wireless Ad hoc Network) เป็นเครือข่ายที่ประกอบขึ้นจากอุปกรณ์คอมพิวเตอร์ที่สามารถเคลื่อนย้ายได้เพื่อเป็นเครือข่ายแบบชั่วคราวและไม่มีการจัดการศูนย์กลาง เครือข่ายไร้สายแบบ Ad hoc นี้สามารถใช้สัญญาณวิทยุจากอุปกรณ์ที่ใช้อยู่ในปัจจุบันนำมาเป็นส่วนเชื่อมต่อกัน ซึ่งข้อดีของเครือข่ายไร้สายแบบ Ad hoc คือ การสร้างทำได้ง่าย สะดวก แต่ในขณะเดียวกันความปลอดภัยของการใช้งานเครือข่ายไร้สายแบบ Ad hoc ก็ถือเป็นสิ่งหนึ่งที่ต้องให้ความสำคัญในการใช้งานด้วย ไม่ว่าจะเป็นการถูกโจมตีในลักษณะต่าง ๆ หรือการถูกขโมยข้อมูล เป็นต้น การรบกวนเครือข่ายไร้สาย ถือเป็นการโจมตีประเภทหนึ่งของ DoS ซึ่งมีรูปแบบการรบกวนหลายประเภท โดยผลของการโจมตีจะทำให้ประสิทธิภาพการทำงานของเครือข่ายลดลง ตามลักษณะการรบกวน

ดังนั้นผู้วิจัยจึงได้ศึกษาเปรียบเทียบการรบกวนเครือข่ายไร้สาย ด้วยรูปแบบการรบกวนแบบพัลส์ และ แบบกวาด เพื่อแสดงให้เห็นถึงประสิทธิภาพของเครือข่ายที่ลดลงจากการถูกรบกวน และค่าต่าง ๆ ของการรบกวนแต่ละประเภทที่มีผลต่อประสิทธิภาพของเครือข่าย เพื่อที่จะใช้เป็นแนวทางสำหรับการวิจัยทางด้านต่าง ๆ เช่น ระบบการตรวจสอบผู้บุกรุก (Intrusion Detection System) หรือ ระบบโต้ตอบผู้บุกรุก (Intrusion Prevention System)

บทความวิจัยนี้ประกอบด้วยส่วนต่าง ๆ ดังนี้ หัวข้อที่ 1 บทนำ หัวข้อที่ 2 วรรณกรรมที่เกี่ยวข้อง ซึ่งได้แก่ เครือข่ายไร้สาย

แบบ Ad hoc การรบกวนเครือข่ายไร้สาย การรบกวนแบบกวาด และการรบกวนแบบพัลส์ หัวข้อที่ 3 วิธีการดำเนินการวิจัย ซึ่ง จะแสดงรายละเอียดของการออกแบบการทดลอง หัวข้อที่ 4 ผลการดำเนินงานจะนำเสนอผลลัพธ์ที่ได้จากการทดลอง และสรุปผลการวิจัยในหัวข้อที่ 5

2. วรรณกรรมที่เกี่ยวข้อง

2.1 เครือข่ายไร้สายแบบ Ad hoc (Wireless Ad hoc

Network)

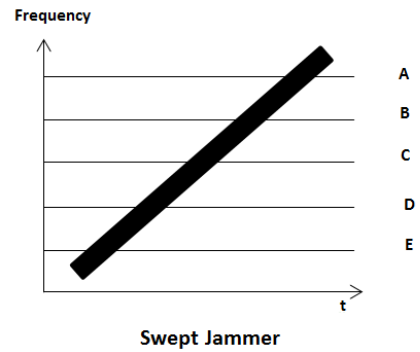
Ad hoc เป็นการสื่อสารข้อมูลระหว่างเครื่องคอมพิวเตอร์ไร้สายและอุปกรณ์ต่าง ๆ ตั้งแต่สองเครื่องขึ้นไป โดยที่ไม่มีศูนย์กลางควบคุม อุปกรณ์ทุกเครื่องสามารถสื่อสารข้อมูลถึงกันได้เอง ซึ่งเครือข่ายไร้สายแบบ Ad hoc นี้สามารถใช้สัญญาณวิทยุจากอุปกรณ์ที่ใช้อยู่ในปัจจุบันนำมาเป็นส่วนเชื่อมต่อกันแต่ละโหนดเป็นได้ทั้งเร้าเตอร์และโฮสต์

2.2 การรบกวนเครือข่ายไร้สาย (Wireless Jamming)

การรบกวนเครือข่ายไร้สาย [2][6] คือ การทำให้อุปกรณ์อิเล็กทรอนิกส์ของฝ่ายตรงข้ามทำงานได้ช้าลง หรือ ทำงานไม่ได้ ซึ่งการรบกวนจะกระทำที่ชั้นกายภาพ (Physical layer) และ ชั้นควบคุมการใช้สื่อ (MAC layer) เนื่องจากง่ายต่อการรบกวนของทั้งสองชั้น เพราะเป็นลักษณะเด่นของเครือข่ายไร้สาย เช่น open medium, signal strength, throughput

2.3 การรบกวนแบบกวาด (Swept Jamming)

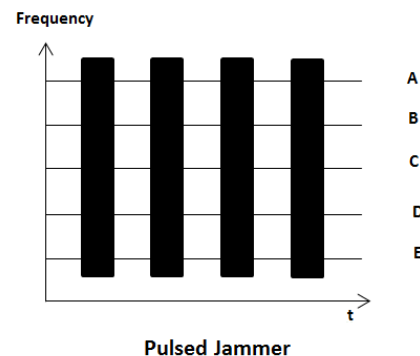
การรบกวนแบบกวาด [2][8] จะทำการรบกวนทีละ 1 ช่องสัญญาณ (Channel) โดยจะทำการกวาดไปที่ละช่องสัญญาณตลอดย่านความถี่ (Bandwidth) ดังภาพที่ 1 ซึ่งการรบกวนแบบนี้จะเป็นการรบกวนแบบอัตโนมัติเพราะมีชุดควบคุมการรบกวน การรบกวนแบบกวาดรวมข้อดีของการรบกวนแบบจุด (Spot Jamming) และการรบกวนแบบเขื่อน (Barrage Jamming) คือ ส่งสัญญาณไปยังช่องสัญญาณแคบ ๆ ตลอดย่านความถี่ ส่วนข้อเสียคือการรบกวนไม่ต่อเนื่อง ถ้าจะให้ต่อเนื่องต้องเพิ่มอัตราความเร็วในการกวาดของสายอากาศจึงจะสามารถรบกวนได้อย่างต่อเนื่อง การรบกวนแบบนี้สามารถตั้งเวลาที่จะหยุดรบกวนแต่ละช่องสัญญาณได้นอกจากนี้การรบกวนแบบกวาดยังง่ายต่อการตรวจจับจากระบบตรวจจับอีกด้วย



ภาพที่ 1: การทำงานของการรบกวนแบบกวาด

2.4 การรบกวนแบบพัลส์ (Pulsed Jamming)

การรบกวนแบบพัลส์ [2][7] จะทำการรบกวนโดยการส่งสัญญาณรูปคลื่นไซน์พัลส์ (Sine-Pulse Waveform) ออกมา โดยการส่งจะส่งออกมาเป็นลักษณะของคาบเวลาที่ขึ้นอยู่กับความเร็วในการหมุน ดังภาพที่ 2 ซึ่งการรบกวนแบบพัลส์จะเป็นการยากในการค้นหาสถานที่ของผู้รบกวน



ภาพที่ 2: การทำงานของการรบกวนแบบพัลส์

2.5 งานวิจัยที่เกี่ยวข้อง

งานวิจัยต่อไปนี้เป็นนำเสนอแนวคิดของผู้วิจัยที่ได้ทำการวิจัยเกี่ยวกับการรบกวนเครือข่ายไร้สาย

Fu และ Chou [3] นำเสนอแบบจำลอง และการจำลองสถานการณ์ของการถูกรบกวนในเครือข่ายไร้สาย โดยได้ศึกษาผลกระทบของการถูกรบกวนแบบต่าง ๆ ทั้งเครือข่ายไร้สายแบบ Infrastructure และ แบบ Ad-hoc ซึ่งผลของการถูกรบกวนก็จะแตกต่างกันออกไป

Yu และคณะ [4] นำเสนอการจำลองและวิเคราะห์ประสิทธิภาพของผู้รบกวนแบบพัลส์ที่มีต่อสภาพแวดล้อมของการสุ่มลำดับสัญญาณรบกวนในระบบ DSSS (Direct Sequence

Spread Spectrum) ซึ่งผลการทดลองแสดงให้เห็นถึงประสิทธิภาพที่ลดลงจากการถูกรบกวน

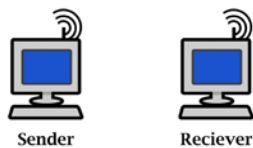
Benslimane และคณะ [5] นำเสนอการวิเคราะห์ผลกระทบต่อการรบกวนบนเครือข่ายไร้สาย 802.11 โดยงานวิจัยนี้ใช้เครื่องมือ NS 3 ในการจำลองสถานการณ์ ซึ่งแสดงค่าทราฟฟิคที่ถูกรบกวนด้วยการโจมตีรูปแบบต่าง ๆ

จากงานวิจัยที่นำเสนอข้างต้นจะแตกต่างจากงานวิจัยฉบับนี้คือ ไม่มีการเปรียบเทียบการโจมตีแบบกวาดกับแบบพัลส์ และนอกจากนั้นงานวิจัยฉบับนี้ได้ทำการศึกษาการรบกวนที่มีต่อเครือข่ายไร้สายแบบ Ad hoc ซึ่งแตกต่างจากที่กล่าวมา

3. วิธีการดำเนินการวิจัย

3.1. ออกแบบการทดลองโดยไม่มีกรรบกวน

ออกแบบเครือข่ายไร้สายแบบ Ad hoc โดยใช้โปรแกรม OPNET 14.5 (Education Version) โดยให้มีโครงสร้างเครือข่ายดังภาพที่ 3



ภาพที่ 3: เครือข่ายไร้สายแบบ Ad hoc เมื่อไม่มีกรรบกวน

ภาพที่ 3 แสดงถึงการออกแบบการทดลองเมื่อไม่มีกรรบกวน เพื่อใช้สำหรับวัดค่าทราฟฟิคของเครื่องผู้รับในขณะที่ไม่ถูกรบกวน และจะนำผลการทดลองไปเปรียบเทียบกับกรรบกวนสัญญาณแบบพัลส์และแบบกวาดต่อไป

โดยการทดลองได้ใช้โปรแกรม OPNET ในการทดลองโดยผู้วิจัยได้สร้างพื้นที่ของการทดลองขนาด 100x100 เมตร เครื่องผู้ส่ง (Sender) จะมีระยะห่างจากเครื่องผู้รับ (Receiver) 10 เมตร โดยกำหนดให้เครื่องผู้ส่งทำการส่งแพ็กเก็ตขนาด 1,024 ไบต์ไปยังเครื่องผู้รับ โดยใช้เวลาในการทดลอง 3 นาที คุณลักษณะที่สำคัญของเครื่องผู้ส่งและเครื่องผู้รับแสดงตามตารางที่ 1

ตารางที่ 1: คุณลักษณะที่สำคัญของเครื่องผู้ส่ง และ ผู้รับ

ผู้ส่ง / ผู้รับ	
Attributes	Values
Physical Characteristics	Direct Sequence
Data Rate	11 Mbps
Frequency	2,426 MHz
Bandwidth	22 MHz
Transmit Power	0.005 W
Packet Size	1,024 Bytes

3.2. ออกแบบการทดลองโดยมีการรบกวนแบบพัลส์



ภาพที่ 4: เครือข่ายไร้สายแบบ Ad-hoc เมื่อมีการรบกวนแบบพัลส์

ภาพที่ 4 ผู้วิจัยได้เพิ่มผู้รบกวนแบบพัลส์ขึ้นมาโดยให้มีระยะห่างจากเครื่องผู้รับ 10 เมตรและสภาพแวดล้อมการทดลองเหมือนขั้นตอนที่ 3.1 หลังจากนั้นก็ปรับคุณลักษณะของผู้รบกวน คือ กำลังตัวส่งของผู้รบกวน เวลาที่ใช้ในการรบกวน และจำนวนผู้รบกวน หลังจากนั้นทำการวัดค่าทราฟฟิคของเครื่องผู้รับเมื่อถูกรบกวนแบบพัลส์ โดยค่าที่วัดได้จะนำไปเปรียบเทียบกับกรรบกวนแบบกวาด โดยคุณลักษณะที่สำคัญของผู้รบกวนแบบพัลส์แสดงตามตารางที่ 2

ตารางที่ 2: คุณลักษณะที่สำคัญของผู้รบกวนแบบพัลส์

ผู้รบกวนแบบพัลส์	
Attributes	Values
Jammer Band Frequency	2,426 MHz
Jammer Bandwidth	22 MHz
Jammer Transmitter Power	0.005 W
Pulsed Width	1 Sec

3.3. ออกแบบการทดลองโดยมีการรบกวนแบบกวาด



ภาพที่ 5: เครื่องข่ายไร้สายแบบ Ad hoc เมื่อมีการรบกวนแบบกวาด

ภาพที่ 5 ผู้วิจัยได้เพิ่มผู้รบกวนแบบกวาดขึ้นมาโดยให้มีระยะห่างจากเครื่องผู้รับ 10 เมตรและสภาพแวดล้อมการทดลองเหมือนขั้นตอนที่ 3.1 หลังจากนั้นก็ปรับคุณลักษณะของผู้รบกวน คือ กำลังตัวส่งของผู้รบกวน เวลาที่ใช้ในการรบกวน และจำนวนผู้รบกวน หลังจากนั้นทำการวัดค่าทราฟฟิคของเครื่องผู้รับเมื่อถูกรบกวนแบบพัลส์ โดยค่าที่วัดได้จะนำไปเปรียบเทียบกับการรบกวนแบบพัลส์ โดยคุณลักษณะที่สำคัญของผู้รบกวนแบบพัลส์แสดงตามตารางที่ 3

ตารางที่ 3: คุณลักษณะที่สำคัญของผู้รบกวนแบบกวาด

ผู้รบกวนแบบกวาด	
Attributes	Values
Jammer Band Base Frequency	2,401 MHz
Jammer Bandwidth	22 MHz
Jammer Transmitter Power	0.005 W
Number of Frequency	60
Cycle Time	120 Sec

3.4. บันทึกผลการทดลอง

หลังจากออกแบบการทดลอง 3 แบบ ขึ้นตอนต่อมาก็คือ บันทึกผลการทดลองทั้งหมด โดยค่าที่เราจะใช้สำหรับการศึกษา และวิเคราะห์ก็คือ ค่าเฉลี่ยทราฟฟิค (Average Throughput) และ ค่าตัวแปร กำลังตัวส่งของผู้รบกวน เวลาที่ใช้ในการรบกวน และจำนวนผู้รบกวน เพื่อดูถึงผลกระทบที่มีต่อค่าเฉลี่ยทราฟฟิคของเครื่องผู้รับ

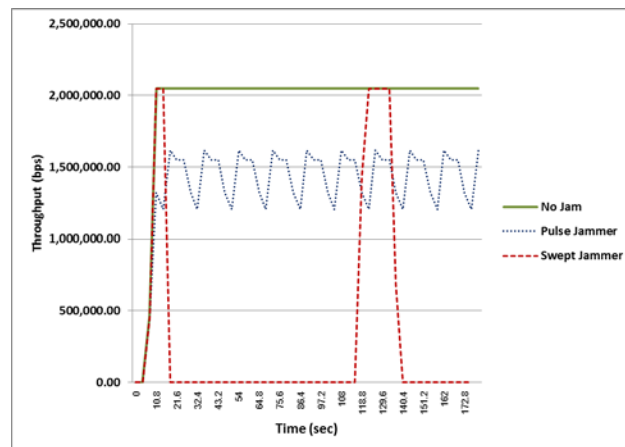
- ค่าเฉลี่ยทราฟฟิค (Average Throughput) คือ จำนวนงานหรือ ข้อมูลที่สามารถส่ง หรือ รับ ได้ในช่วงเวลาหนึ่ง ซึ่งสามารถหาได้จาก (1)

$$Average\ Throughput = Data / Time \quad (1)$$

4. ผลการดำเนินงาน

4.1 เปรียบเทียบค่าเฉลี่ยทราฟฟิค

ผลการเปรียบเทียบค่าทราฟฟิคของเครื่องผู้รับในสภาพแวดล้อมที่ไม่มีมีการรบกวน (การทดลองที่ 3.1) ถูกรบกวนแบบพัลส์ (การทดลองที่ 3.2) และ ถูกรบกวนแบบกวาด (การทดลองที่ 3.3) โดยใช้คุณลักษณะที่กำหนดในตารางที่ 1-3 แสดงดังภาพที่ 6



ภาพที่ 6: เปรียบเทียบค่าทราฟฟิคที่เครื่องผู้รับ เมื่อกำลังตัวส่งของผู้รบกวนเท่ากับ 0.005 W

ภาพที่ 6 เมื่อไม่มีมีการรบกวนค่าเฉลี่ยทราฟฟิคของเครื่องผู้รับจะมีค่า 1.93 Mbps เมื่อถูกรบกวนแบบพัลส์ค่าเฉลี่ยทราฟฟิคจะลดลงเหลือ 1.36 Mbps ลดลง 17 % จากเมื่อไม่มีมีการรบกวน และค่าเฉลี่ยทราฟฟิคเมื่อถูกรบกวนแบบกวาดจะมีค่า 0.29 Mbps ลดลง 85 % จากเมื่อไม่มีมีการรบกวน ซึ่งผลของการเปรียบเทียบแสดงให้เห็นว่าการรบกวนแบบกวาดทำให้ค่าเฉลี่ยทราฟฟิคของเครื่องผู้รับลดลงมากกว่าการรบกวนแบบพัลส์

4.2 เปรียบเทียบค่าทราฟฟิคเมื่อกำลังตัวส่งของผู้รบกวนต่างกัน

ทดลองปรับกำลังตัวส่งของผู้รบกวนให้ต่างกันจากการถูกรบกวนแบบพัลส์ (การทดลองที่ 3.2) และ ถูกรบกวนแบบกวาด (การทดลองที่ 3.3) โดยให้มีค่า 0.01, 0.005, 0.1 และ 1 วัตต์ ผลจากการทดลองแสดงดังตารางที่ 4

ตารางที่ 4: เปรียบเทียบค่าเฉลี่ยทราฟฟิกของเครื่องผู้รับเมื่อพลังงานของผู้รับต่างกัน

ค่าเฉลี่ยทราฟฟิกของเครื่องผู้รับ (Mbps)				
พลังงาน (W)	0.001	0.005	0.1	1
ชนิด				
Pulse Jammer	1.36	1.36	1.36	1.36
Swept Jammer	0.29	0.29	0.29	0.29

ตารางที่ 4 แสดงให้เห็นว่าค่าพลังงานของผู้รับที่ต่างกันไม่ทำให้ค่าเฉลี่ยทราฟฟิกของการรบกวนแบบพัลส์และการรบกวนแบบกวาดแตกต่างกัน

4.3 เปรียบเทียบค่าทราฟฟิกเมื่อระยะเวลาที่ใช้ในการรบกวนต่างกัน

ทดลองปรับระยะเวลาที่ใช้ในการรบกวนให้ต่างกันจากการรบกวนแบบพัลส์ (การทดลองที่ 3.2) และ การรบกวนแบบกวาด (การทดลองที่ 3.3) โดยให้มีค่า 1, 3, 7, 10 นาที ซึ่งเลือกใช้ระยะเวลาในการรบกวนสั้น ปานกลาง และระยะยาว ซึ่งผลจากการทดลองแสดงดังตารางที่ 5

ตารางที่ 5: เปรียบเทียบค่าเฉลี่ยทราฟฟิกของเครื่องผู้รับเมื่อระยะเวลาที่ใช้ในการรบกวนต่างกัน

ค่าเฉลี่ยทราฟฟิกของเครื่องผู้รับ (Mbps)				
เวลา (นาที)	1	3	7	10
ชนิด				
Pulse Jammer	1.19	1.36	1.41	1.42
Swept Jammer	0.27	0.29	0.29	0.27

ตารางที่ 5 แสดงให้เห็นว่าเวลาที่ใช้ในการรบกวนของผู้รบกวนแบบพัลส์และแบบกวาด ให้ค่าเฉลี่ยทราฟฟิกของเครื่องผู้รับแตกต่างกัน โดยที่การรบกวนแบบพัลส์นั้นค่าเฉลี่ยทราฟฟิกของเครื่องผู้รับนั้นจะน้อยในตอนแรก ๆ และจะสูงขึ้นในระยะเวลาที่เพิ่มขึ้นเนื่องจากคุณสมบัติของการรบกวนแบบพัลส์ที่จะมีประสิทธิภาพในการรบกวนตอนระยะเวลาแรก ๆ ของการโจมตี หลังจากนั้นประสิทธิภาพในการรบกวนจะลดลง

การรบกวนแบบกวาดนั้นให้ค่าเฉลี่ยทราฟฟิกของเครื่องรบกวนต่างกันในแต่ละนาที เนื่องจากคุณสมบัติของการรบกวนแบบกวาดนั้นจะกวาดสัญญาณไปที่ละช่องสัญญาณซึ่งระยะเวลาของการรบกวนนั้นจะถูกรบกวนไม่เท่ากัน

4.4 เปรียบเทียบค่าทราฟฟิกเมื่อจำนวนผู้รับต่างกัน

ทดลองปรับจำนวนของผู้รับกวนให้ต่างกันจากการรบกวนแบบพัลส์ (การทดลองที่ 3.2) และ การรบกวนแบบกวาด (การทดลองที่ 3.3) โดยให้มีจำนวนของผู้รับกวน 1, 2, 5 และ 10 ซึ่งผลจากการทดลองแสดงดังตารางที่ 6

ตารางที่ 6: เปรียบเทียบค่าเฉลี่ยทราฟฟิกของเครื่องผู้รับเมื่อจำนวนของผู้รบกวนต่างกัน

ค่าเฉลี่ยทราฟฟิกของเครื่องผู้รับ (Mbps)				
จำนวน	1	2	5	10
ชนิด				
Pulse Jammer	1.36	1.36	1.36	1.36
Swept Jammer	0.29	0.29	0.29	0.29

ตารางที่ 6 แสดงให้เห็นว่าเมื่อเพิ่มผู้รบกวนของการโจมตีแบบพัลส์และแบบกวาดเพิ่มขึ้น ค่าเฉลี่ยทราฟฟิกของเครื่องผู้รับไม่เกิดการเปลี่ยนแปลง ซึ่งจะกล่าวได้ว่าจำนวนของผู้รบกวนไม่ส่งผลต่อค่าเฉลี่ยทราฟฟิกของเครื่องผู้รับ

5. สรุปผลการวิจัย

จากการศึกษาผลกระทบของการรบกวนเครือข่ายไร้สายแบบพัลส์และแบบกวาดที่มีต่อเครือข่ายไร้สายแบบ Ad Hoc สามารถสรุปได้ดังนี้

5.1 การรบกวนแบบกวาดส่งผลต่อค่าเฉลี่ยทราฟฟิกมากกว่าการรบกวนแบบพัลส์ ซึ่งการรบกวนแบบกวาดจะทำให้ค่าเฉลี่ยทราฟฟิกของเครื่องผู้รับลดลง 85 % เมื่อเปรียบเทียบกับในขณะที่ไม่ถูกรบกวน ส่วน การรบกวนแบบพัลส์จะทำให้ค่าเฉลี่ยทราฟฟิกของเครื่องผู้รับลดลง 17 % เมื่อเปรียบเทียบกับในขณะที่ไม่ถูกรบกวน

5.2 กำลังตัวส่งของผู้รบกวน ของการรบกวนแบบพัลส์ และแบบกวาด ไม่ส่งผลต่อค่าเฉลี่ยทราฟฟิกของเครื่องที่ถูกรบกวน

5.3 เวลาที่ใช้ในการรบกวนของผู้รบกวนแบบกวาดจะมีผลต่อค่าเฉลี่ยทราฟฟิกของเครื่องผู้รับมากกว่าการรบกวนแบบพัลส์ เนื่องจากการรบกวนแบบกวาดนั้นคุณลักษณะการกวาดช่องสัญญาณขึ้นอยู่กับจำนวนเวลา

5.4 จำนวนของผู้รบกวนแบบพัลส์และกวาดที่เพิ่มขึ้นไม่ส่งผลต่อค่าเฉลี่ยทราฟฟิกของเครื่องผู้รับ

จากผลการวิจัยเราสามารถนำผลที่ได้ไปประยุกต์ใช้กับการพัฒนาระบบการตรวจสอบผู้บุกรุก หรือ ระบบโต้ตอบผู้บุกรุกโดยใช้ตัวแปรที่มีผลต่อการรบกวนแต่ละประเภทในการพัฒนาระบบให้มีประสิทธิภาพเพิ่มขึ้น

เอกสารอ้างอิง

- [1] “Wireless LAN”, (เอกสารออนไลน์) สารสังเขป สืบค้นวันที่ 9 ธันวาคม จาก <http://center.finansalife.com/qagt/wirelesslan.htm>
- [2] น.อ.อำนาจ ทองรอ, “การประมวลผลสัญญาณของอุปกรณ์สงครามอิเล็กทรอนิกส์”, จดสาร สวพ.ทร.
- [3] Tian Fu and Te-Shun Chou, “Modeling and Simulation of Jamming Attacks in WLAN”, *Master of Thesis*, April, 2012.
- [4] Zhang Yu, Xuehe Zheng, Sen Yang and Ning Zhong , “Performance analysis and simulation of DSSS system in classical urban fading environments and pulse jammer”, Information Science and Engineering (ICISE), 2010 2nd International Conference, 2010.
- [5] Benslimane A, El yakoub and Bouhorma M, “Analysis of Jamming effect on IEEE 802.11 Wireless Networks”, Communications (ICC), 2011 IEEE International Conference on 2011.
- [6] D. J. Thuente and M. Acharya, “Intelligent Jamming in Wireless Networks with Applications to 802.11b and Other Networks”, MILCOM '06, 2006.
- [7] Bayraktaroglu E., King C., Liu X., Noubir G., Rajaraman R., Thapa, B.,” On the Performance of IEEE 802.11 under Jamming”, INFOCOM 2008, 2008.
- [8] Benslimane, A., El yakoubi, A. and Bouhorma M., "Analysis of Jamming Effects on IEEE 802.11 Wireless Networks", Communications (ICC), 2011 IEEE International Conference, 2011.

การป้องกันการโจมตี SQL injection โดยการเข้ารหัสแบบสมมาตรสำหรับระบบ

อินทราเน็ต

SQL Injection Attack Prevention by Symmetric Cryptography for Intranet Application

ชุตินา อำไพรัตน์ (Chutima Ampairat)¹ และศิริปรัชญ์ บุญครอง (Sirapat Boonkrong)²

^{1,2}ภาควิชาการสื่อสารข้อมูลและเครือข่าย คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹ampairat_5@hotmail.com, ²sirapatb@kmutnb.ac.th

บทคัดย่อ

การโจมตีในรูปแบบของ SQL injection ถือเป็น การโจมตีที่ยังได้รับความนิยมเป็นอันดับหนึ่งในปี 2013 [1] วิธีนี้สามารถเข้าถึงการจัดการฐานข้อมูลทั้งหมดได้ โดยการโจมตีแบบ SQL injection จะทำการตรวจสอบหาช่องโหว่ในการใส่ค่าในส่วนที่เป็น user interface ของ เว็บไซต์ จากนั้นผู้โจมตีก็จะสามารถเปลี่ยนแปลงบางส่วนของ web request เพื่อให้สามารถ query ข้อมูลจาก ฐานข้อมูลได้จึงส่งผลให้ข้อมูลที่เป็นความลับของผู้ใช้ไม่ปลอดภัยอีกต่อไปงานวิจัยนี้จึงมีแนวคิดเพื่อป้องกันการโจมตี SQL injection เพื่อให้รักษาความลับของข้อมูลและมีความปลอดภัยในการใช้งานระบบมากยิ่งขึ้น

คำสำคัญ: เอสคิวแอลอินเจกชัน การเข้ารหัส การถอดรหัส

Abstract

SQL injection has been one of the most used attacks in the past years. The attack allows attacker to gain access into a web site's database. It is done by an attacker inserting an input via the web site's user interface. The attacker is then able to make changes to a web request that, in turn, will help the attacker to query some data from the database. This research, therefore, designs and develops a method for preventing SQL injection so that the data on a website is more secure.

Keyword: SQL Injection, Encryption, Decryption

1. บทนำ

ปัจจุบันการใช้บริการ web service ถือว่าสะดวกในการทำธุรกรรมต่างๆเป็นอย่างมาก [2] และมีการใช้กันอย่างแพร่หลายในสังคมออนไลน์ ซึ่งในการใช้งานในแต่ละเว็บไซต์นั้น เว็บเซิร์ฟเวอร์ก็จะมีเก็บข้อมูลการใช้งานของผู้ใช้ไว้ฐานข้อมูลเพื่อความสะดวก และรวดเร็ว ในการใช้งานครั้งต่อไป แต่อย่างไรก็ตาม นอกจาก web service เหล่านี้จะมีข้อดีแล้ว ก็ยังมีช่องโหว่ ที่ทำให้ผู้โจมตีสามารถใช้วิธีการต่างๆ เพื่อขโมยข้อมูลที่เป็นความลับหรือแก้ไขข้อมูลทำให้เกิดความเสียหายได้ การโจมตีในรูปแบบ SQL injection มีการป้องกันที่หลากหลายวิธีไม่ว่าจะเป็นการป้องกันจากเว็บเซิร์ฟเวอร์ หรือ การตรวจจับด้วย application ต่างๆ และ Encryption Algorithm ที่ค่อนข้างซับซ้อนพอสมควร การเสนอแนวคิดและวิธีการต่างๆก็ยังคงมีการทดสอบอย่างต่อเนื่องเพื่อหาวิธีการตรวจจับและป้องกันที่ดีกว่า การใช้ signature ในการป้องกันเช่นการใช้อุปกรณ์ต่างๆ เช่น Firewall , IPS ที่จะคอยตรวจดู traffic ใน Network ที่จะคอยตรวจจับข้อความที่แปลกปลอมใน packet ที่เหมือนกับ signature [3] ก็ยังพบว่ามีช่องโหว่ที่สามารถโจมตีได้ เพราะผู้โจมตียังมีการพัฒนารูปแบบการโจมตีอยู่ตลอดเวลา ความปลอดภัยของข้อมูลจึงเป็นสิ่งสำคัญต่อการดำเนินการ หากเกิดความเสียหายก็จะส่งผลกระทบต่อ ไม่ว่าจะมากหรือน้อยก็ จะทำให้องค์กรขาดความมั่นใจได้ [4] ดังนั้นจึงต้องมีการรักษาความปลอดภัยของข้อมูลเพื่อไม่ให้ข้อมูลนั้นถูกกระทำความเสียหายและถูกคุกคามไปในทางที่ไม่ถูกต้องได้

งานวิจัยนี้จึงมีวัตถุประสงค์มีวัตถุประสงค์เพื่อออกแบบและพัฒนาวิธีการป้องกันการโจมตีแบบ SQL Injection โดยประยุกต์ใช้การเข้ารหัสข้อมูลสำหรับเว็บไซต์ในองค์กรเพื่อ

ไม่ให้เกิดความเสียหายได้ ซึ่งงานวิจัยนี้ได้จัดหัวข้อตามลำดับ ดังนี้ คือ ส่วนที่ 2 อธิบายถึงทฤษฎีและงานวิจัยที่เกี่ยวข้อง ส่วนที่ 3 วิธีดำเนินการวิจัย ส่วนที่ 4 ผลการดำเนินงาน และบทสรุปของงานวิจัยในส่วนที่ 5

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การโจมตีรูปแบบ SQL injection

การโจมตีแบบ SQL Injection คือการที่ผู้โจมตี ใช้ช่องโหว่ของโปรแกรมทำให้สามารถเติมคำสั่งภาษา SQL เข้าไปทางส่วนรับ ของโปรแกรมได้ [5,6] หากเว็บแอปพลิเคชันที่เปิดช่องให้ผู้โจมตีทำการโจมตีแบบนี้ได้ ก็จะทำให้แฮกเกอร์สามารถอ่านข้อมูลสำคัญต่างๆ ในระบบฐานข้อมูล (SELECT) แก้ไขข้อมูลต่าง ๆ ในระบบฐานข้อมูล

(INSERT/UPDATE/DELETE) เรียกคำสั่งที่ใช้ในการจัดการระบบฐานข้อมูลเช่นคำสั่งสำรองข้อมูล (backup)หรือแม้กระทั่งเรียกคำสั่งของระบบปฏิบัติการได้ในบางกรณีซึ่งSQL Injection จะเกิดขึ้นได้ต้องมีปัจจัยสองประการ

โปรแกรมรับข้อมูลจากแหล่งที่ไม่น่าเชื่อถือต่างๆ เช่น Form Input, query String, Cookie เป็นต้น ซึ่งข้อมูลดังกล่าวไม่มีการตรวจสอบความถูกต้องของข้อมูลก่อน ยกตัวอย่าง เช่น หากข้อมูลที่รับเข้ามานั้นคือหมายเลขโทรศัพท์ ก็ไม่ได้ตรวจสอบว่าข้อมูลดังกล่าวจะต้องประกอบด้วยตัวเลขเท่านั้น หรือ ไม่ได้ตรวจสอบว่าข้อมูลดังกล่าวมีตัวอักษรอันตรายที่ใช้ในภาษา SQL เช่น single quote ('), dash (-), back slash (\) หรือไม่ เป็นต้น

2.2 การรักษาความปลอดภัยของข้อมูล

เทคนิคและวิธีการรักษาความปลอดภัยของข้อมูลสามารถทำได้หลายวิธีเพื่อเป็นการเก็บข้อมูลขององค์กรไว้เป็นความลับ และป้องกันผู้ที่ไม่ประสงค์ดีทำการเปลี่ยนแปลงข้อมูลซึ่งส่งผลให้เกิดความเสียหายในเชิงธุรกิจได้ [7]

การเขียนโปรแกรมให้รัดกุมมากขึ้น โดยเลือกใช้เครื่องมือต่างๆ ที่แต่ละภาษามีไว้ให้เพื่อช่วยในการคัดกรองความผิดปกติ เช่น คำสั่ง mysql_real_escape_string() สามารถใช้กับภาษา PHP ที่ทำงานร่วมกับ MySQL Database จะช่วยคัดกรองอักขระต่างๆ ที่ต่อเค้านการละเมิดได้ ฉะนั้นก่อนนำค่าที่ส่งมาจากผู้ใช้ไปใช้งาน ควรนำมาผ่านขั้นตอนนี้ก่อน เป็นต้น

```
$username=mysql_real_escape_string($_POST['username']);
$password=mysql_real_escape_string($_POST['password']);
```

การกรองข้อมูลจะต้องทำทุกครั้งก่อนนำไปใช้งาน ตัวอย่างต้นทางที่ควรต้องกรองเช่น Form Input ,users, parameters URL, Values from Cookie

กรณีข้อมูลที่เป็นตัวเลข ควรทำการตรวจสอบ โดยการแปลงค่าให้เป็นตัวเลข หรือใช้ฟังก์ชัน ISNUMERIC เพื่อตรวจสอบการเป็นตัวเลขจริงๆ

ดูแลเกี่ยวกับสิทธิของผู้ใช้ที่สามารถ เข้าสู่ ฐานข้อมูล ให้อยู่ในวงจำกัด ไม่เปิดมากเกินไป และไม่ควรถือ user ที่มีสิทธิ์เป็น root มาใช้งาน เพราะถ้าหากเกิดช่องโหว่แล้วผู้ไม่หวังดีสามารถใช้สิทธิ์ของ user นั้นทำความเสียหายต่อระบบมากขึ้น

กรองตัวอักขระที่เข้าข่ายการละเมิดออกจากข้อมูลอินพุท เช่น double quote, single quote, slash, back slash, semi colon, extended character like NULL, carry return, new line และอื่นๆ ที่ไม่ควรจะเป็นข้อมูลที่ใช้ในระบบ [8]

ควรกำหนดและตรวจสอบความยาวของข้อมูลที่ส่งมาทุกครั้งเพื่อป้องกันการแฝงคำสั่งอื่นๆ เข้ามาในฟิลด์ข้อมูลปกติ

ใช้เครื่องมือช่วยตรวจจับและกรองข้อมูล เช่น โปรแกรมประเภทเว็บแอปพลิเคชันไฟ หรือระบบ IPS (Intrusion Prevention System) วิธีการนี้อาจต้องแลกด้วยประสิทธิภาพที่ลดลงเนื่องมาจากการดักจับข้อมูลและการตรวจสอบ

2.3 การเข้ารหัสข้อมูล

การเข้ารหัสแบบสมมาตร (symmetric cryptography) เป็นกระบวนการสำหรับการแปรรูปข้อมูลเล็กทรอนิกส์ธรรมดาให้อยู่ในรูปแบบที่บุคคลทั่วไปไม่สามารถอ่านเข้าใจได้ ซึ่งโดยทั่วไปแล้วการเข้ารหัสจะกระทำก่อนการจัดเก็บข้อมูลหรือก่อนการส่งข้อมูล โดยการนำข้อมูลอิเล็กทรอนิกส์ธรรมดากับกุญแจ (คีย์) ซึ่งเป็นตัวเลขสุ่มใดๆ มาผ่านกระบวนการทางคณิตศาสตร์ ผลที่ได้ก็คือข้อมูลที่เข้ารหัส ขั้นตอนที่กำลังจะกล่าวมานี้จะเรียกว่า “การเข้ารหัส” (encryption) และเมื่อต้องการอ่านข้อมูล ก็นำเอาข้อมูลที่เข้ารหัสกับกุญแจมาผ่านกระบวนการทางคณิตศาสตร์ ผลลัพธ์ที่ได้ก็คือข้อมูลดั้งเดิม ซึ่งขั้นตอนนี้จะเรียกว่า “การถอดรหัส” (decryption) งานวิจัยนี้จึงทำการพัฒนา อัลกอริทึมเพื่อนำมาประยุกต์ใช้ในการป้องกันการโจมตีผ่านเว็บไซต์

(Website) ภายในองค์กร โดยการเข้ารหัส (encryption) เพื่อให้สามารถรักษาข้อมูลให้ปลอดภัย

อัลกอริทึม AES [9] เป็นการเข้ารหัสแบบสมมาตรซึ่งจะเข้ารหัสลับข้อมูลเป็นกลุ่ม กลุ่มละ 128 บิต กฎเกณฑ์ที่ใช้ในการเข้าและถอดรหัสมีขนาด 128 192 หรือ 256 บิต ขึ้นอยู่กับผู้ใช้งาน ขนาดของกุญแจจะเป็นตัวกำหนดจำนวนรอบของการประมวลผลซึ่งอาจแสดงในรูปของบิตหรือค่าก็ได้โดยกุญแจขนาด n บิต จะมีขนาดความยาวของกุญแจ (Nk) เท่ากับ $n/32$ ค่าขนาดของกุญแจจะเป็นตัวกำหนดจำนวนรอบของการประมวลผล (Nr)

อัลกอริทึม AES ถือเป็นอัลกอริทึมแบบสมมาตร ที่ได้รับการยอมรับจากองค์กรต่าง ๆ และได้รับความนิยมจากคนทั่วโลกมากที่สุด ทั้งนี้เนื่องมาจากความรวดเร็วของการประมวลผลในกระบวนการเข้าและถอดรหัสลับ โดยได้รับการออกแบบให้มีการทำงานที่เหมาะสมกับ โพรเซสเซอร์ (Processor) รุ่นใหม่ ๆ ความต้องการหน่วยความจำเพื่อใช้ในการเข้าและถอดรหัสลับที่น้อย รวมทั้งความง่ายในการนำไปประยุกต์ใช้งานทั้งในรูปของซอฟต์แวร์และฮาร์ดแวร์ได้อย่างกว้างขวาง

งานวิจัยนี้จึงได้นำหลักการเข้ารหัสแบบสมมาตร โดยเลือกอัลกอริทึม AES 128 เพื่อมาช่วยในการรักษาความปลอดภัยของข้อมูลในการใช้งานบนเว็บแอปพลิเคชันเนื่องจากการเข้ารหัสข้อมูล จะทำให้ค่าที่ได้รับมีขนาดและรูปแบบที่เปลี่ยนไปจึงทำให้ยากต่อการคาดเดาและการโจมตีมากยิ่งขึ้น

2.4 งานวิจัยที่เกี่ยวข้อง

การป้องกัน SQL injection โดยส่วนใหญ่จะเป็นการใช้ Tools ในการป้องกัน [10] และนอกจากนี้ยังมีนักวิจัยคิดค้นเกี่ยวกับการสร้างกฎไวยากรณ์ภาษาเอสคิวแอล สำหรับการวิเคราะห์คำสั่งเอสคิวแอล อินเจกชัน โดยประยุกต์จากรูปแบบมาตรฐาน Bakus Naur Form (BNF) เพื่อตรวจสอบค่าที่ได้รับจากผู้ใช้แล้วนำมาเปรียบเทียบกับค่าตามกฎของไวยากรณ์หากมีค่าที่มีการเรียกใช้งานเหมือนกับค่าตามกฎของไวยากรณ์ ผู้ใช้ก็จะสามารถเข้าใช้ระบบได้ [11] ตัวอย่างกฎไวยากรณ์สำหรับตรวจสอบสิทธิ์การเข้าใช้ระบบ

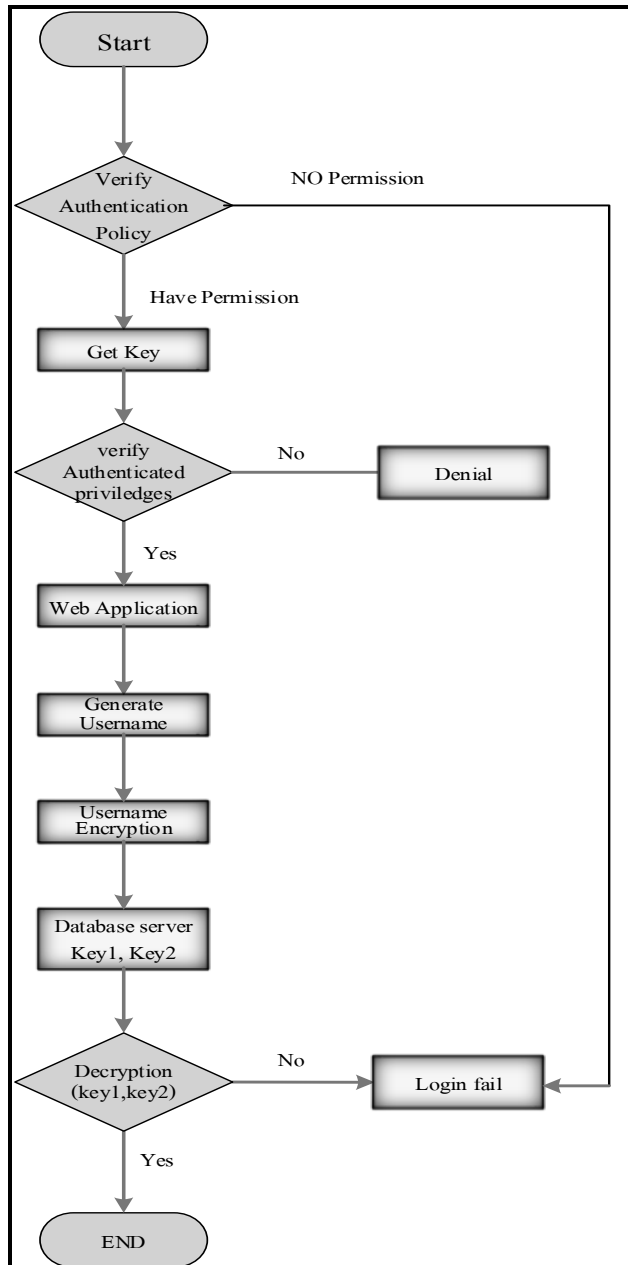
ภาพที่ 1: ตัวอย่างกฎไวยากรณ์สำหรับตรวจสอบสิทธิ์การเข้าใช้ระบบ

นอกจากการนำระบบวิเคราะห์ทางไวยากรณ์มาใช้งานแล้ว ยังนำการเข้ารหัสเพื่อเข้ารหัสข้อมูล และมีการเก็บข้อมูลการใช้งาน โดยใช้ MD5 hashing [12] วิธีนี้จะช่วยลดขนาดของการเก็บข้อมูลวิธีการติดตั้ง โปรแกรม Mod Security [13] เพื่อป้องกันการโจมตีจาก SQL injection โดยการทดลองได้นำ signature ของ Mod Security มาเป็นนโยบายของระบบป้องกันภัยซึ่งเป็นเว็บแอปพลิเคชันไฟร์วอลล์ ติดตั้งโดยใช้ SMF web Broad และเขียน PHP สคริป ในการป้องกันการขโมย Session ID และกำหนดชนิดไฟล์และขนาดของไฟล์ที่จะใช้อัปเดตเว็บแอปพลิเคชัน นอกจากนั้นยังทำฐานข้อมูลเพื่อเก็บข้อมูลการใช้งานต่างๆ ของผู้ใช้งานเว็บแอปพลิเคชัน [14] การตรวจสอบการโจมตีโดยใช้ชื่อว่า K-centers [15] ซึ่งได้เสนอการตรวจสอบการโจมตี SQL injection คือตรวจสอบตัวเลข และ Centers ในกลุ่มของ KC เมื่อได้รับค่าที่เราไม่สามารถมองเห็น และไม่สามารถตรวจสอบค่าที่แตกต่างกันได้วิธีการนี้ไม่ได้ส่งผลต่อค่าที่ปกติ และผิดปกติ เมื่อได้รับค่าที่แตกต่างกันจากการโจมตีแต่อย่างไรก็ตามยังมีข้อเสียเปรียบเนื่องจากว่าต้องมีการรับ label ที่แท้จริงของแต่ละคำสั่งเท่านั้น

จากที่ได้ศึกษางานวิจัยที่เกี่ยวข้องจะแตกต่างจากงานวิจัยนี้ ในส่วนของการเข้ารหัสไปใช้งานบนอินเทอร์เน็ตซึ่งเป็นเว็บแอปพลิเคชันภายในองค์กร

3. วิธีดำเนินการวิจัย

ในงานวิจัยนี้ได้ออกแบบวิธีการเข้ารหัส query เพื่อให้ปลอดภัยจากการโจมตีและรักษาความลับข้อมูลขององค์กร ซึ่งจะประกอบไปด้วย 3 ส่วนการทำงานหลักคือ การ Generate อินพุต การเข้ารหัสข้อมูล และการทดสอบการโจมตีด้วยการใช้ค่าพารามิเตอร์



ภาพที่ 2: โครงสร้างระบบ

จากภาพที่ 2 ผู้ใช้จะต้องผ่านการตรวจสอบสิทธิ์ในการใช้งานโดยผ่านการตรวจสอบจาก Policy ขององค์กร โดยผู้ใช้งานจะต้องเป็นพนักงานเท่านั้นจึงจะมีสิทธิ์เข้าสู่ระบบภายในขององค์กรได้เมื่อผู้ใช้งานผ่านการตรวจสอบสิทธิ์เรียบร้อยแล้ว หาก

เป็นพนักงานก็จะมีรหัสพนักงานจำนวน 6 หลักเท่านั้นที่จะยืนยันตัวตนในการเข้าใช้ระบบได้พร้อมกับผู้ใช้จะถือคีย์ที่ตรงกับทางเซิร์ฟเวอร์มีไว้เพื่อสามารถเข้าและถอดรหัสได้ ซึ่งการเข้าใช้ระบบจะต้องทำการล็อกอินโดยใช้รหัสพนักงาน

หลังจากการยืนยันตัวตนถูกต้องตามขั้นตอนเรียบร้อยแล้ว ผู้ใช้จะสามารถเข้าสู่ระบบต่างๆของภายในหน่วยงานได้เริ่มจากการล็อกอินระบบ โดยจะมีขั้นตอนในการเข้าระบบคือเมื่อผู้ใช้ป้อนอินพุตระบบก็จะทำการ Generate อินพุตเพื่อแปลงให้อยู่ในรูปแบบของ Ciphertext แล้วเข้าสู่การเข้ารหัสโดยใช้ อัลกอริทึม AES 128 โดยค่าที่ได้จะถูกส่งไปยังเซิร์ฟเวอร์ เป็นการแปลงข้อมูลให้อินพุตที่ได้รับมีความซับซ้อนมากยิ่งขึ้น เมื่อทางเซิร์ฟเวอร์ได้รับ Ciphertext เรียบร้อยแล้วก็จะทำการถอดรหัสข้อมูลโดยใช้คีย์ตัวเดียวกับทางฝั่งผู้ใช้ หากคีย์ที่ผู้ใช้ส่งมานั้นตรงกันกับทางฝั่งเซิร์ฟเวอร์ ผู้ใช้ก็จะสามารถเข้าสู่ระบบต่างๆขององค์กรได้ในกรณีที่คีย์ที่ใช้ในการเข้าและถอดรหัสของผู้ใช้และเซิร์ฟเวอร์ไม่ตรงกัน ผู้ใช้จะไม่สามารถเข้าสู่ระบบได้

ระบบจะสามารถป้องกัน SQL injection ได้จากการเข้ารหัสข้อมูลเนื่องจากการเข้ารหัสนั้นจะสามารถเรียกใช้งานข้อมูลได้เฉพาะผู้ที่มีคีย์ตัวเดียวกันกับเซิร์ฟเวอร์ เท่านั้นจึงทำให้ระบบมีความปลอดภัยในการใช้งานและการตรวจสอบผู้บุกรุก

การเปรียบเทียบและวิเคราะห์ ลบค่า Attribute ของ SQL query ระบบจะทำการ scan fixed SQL query ในการใช้งานเว็บ และรายงานผลการเปรียบเทียบเวลาในการใช้งานจริงเพื่อสร้าง Dynamic SQL query จากอินพุตของผู้ใช้งานโดยจะถูกสร้างขึ้นจากผู้ใช้งานปกติ ระหว่าง เว็บแอปพลิเคชัน และฐานข้อมูลเซิร์ฟเวอร์ แล้วนำมาเปรียบเทียบกับการสร้างแบบ dynamic และนำมาตรวจจับการบุกรุก SQL injection

งานวิจัยนี้ได้ทำการทดสอบประสิทธิภาพของระบบโดยการจำลองเว็บแอปพลิเคชันและเข้ารหัสข้อมูลที่ผู้ใช้ได้ทำการเข้าระบบผ่านหน้าเว็บอินเตอร์เฟส

3.1 การ Generate input

ในขั้นตอนนี้จะเป็นการ Generate อินพุตที่ผู้ใช้ป้อนเข้าไปเมื่อต้องการเข้าสู่ระบบโดยจะมีการ generate โดยใช้ machine คีย์ ซึ่งจะเป็นการแปลงค่าอินพุต ที่ได้รับจากหน้าอินเตอร์เฟสแล้วนำมา generate เพื่อให้ได้อินพุต ที่ไม่เป็นตัวอักษรที่ได้

รับมาจากผู้ใช้ซึ่งในการ generate แต่ละครั้งอินพุตที่ได้รับจะมีค่าที่ไม่ซ้ำกัน

แล้ว ระบบก็จะทำการถอดรหัส ที่อยู่ในรูปของ MachineKey อีกครั้งเพื่อความปลอดภัยสำหรับผู้ใช้งาน

ภาพที่ 3: การ Generate อินพุต

3.2 การเข้ารหัสข้อมูล

หลังจากที่ได้รับอินพุตที่ผ่านการ Generate จาก machineKey เรียบร้อยแล้วก็จะนำค่าอินพุตที่ได้รับมาเข้ากระบวนการเข้ารหัสด้วยอัลกอริทึม AES 128 ก่อนที่จะส่งไปยังเซิร์ฟเวอร์ เพื่อ query ข้อมูลในฐานข้อมูลเพื่อใช้งานในระบบ ซึ่งจะเป็นการป้องกันความปลอดภัยเป็น 2 ชั้น



ภาพที่ 4: การเข้ารหัสด้วยอัลกอริทึม AES 128

ภาพที่ 6: การถอดรหัสในส่วนของ machineคีย์

4. ผลการดำเนินงาน

งานวิจัยนี้ได้ดำเนินการเข้ารหัสข้อมูล เพื่อการรักษาความปลอดภัยของข้อมูล โดยทำการ สร้าง คีย์ และ เข้ารหัสของมูลถึง 2 ครั้งเพื่อที่จะเข้าใช้งานระบบ ผู้วิจัยจึงทำการทดสอบด้านความปลอดภัยโดยประเมินผลจากการจำลองการโจมตีของระบบด้วยการใช้ค่าพารามิเตอร์ซึ่งมีผลดังนี้

3.3 การทดสอบการโจมตีด้วยการใส่ค่าพารามิเตอร์

ทำการทดสอบการใช้งานระบบโดยการจำลอง web login โดยใช้ค่าพารามิเตอร์เพื่อทำการเข้าสู่ระบบที่ยังไม่มีการใช้หลักการในการเข้ารหัสข้อมูลซึ่งจะใช้ค่าพารามิเตอร์คือ "anything' or 'x'='x'" แล้วทำการเข้าสู่ระบบ

```
SELECT fieldlist
FROM table
WHERE field = 'anything' OR 'x'='x';
```

ภาพที่ 5: ค่าพารามิเตอร์ที่ใช้ในการทดสอบระบบ

ภาพที่ 7: เข้าสู่ระบบโดยใช้ค่าพารามิเตอร์

เมื่อทำการ Login เข้าสู่ระบบโดยใช้พารามิเตอร์ ผู้โจมตีจะสามารถเข้าสู่ระบบไปยังฐานข้อมูลได้ทันทีจึงส่งผลให้ข้อมูลเกิดความเสียหายได้ง่าย หลังจากใช้อัลกอริทึม AES 128 เข้ารหัส ข้อมูลแล้วทำการทดสอบได้ผลดังนี้

ภาพที่ 8: เข้าสู่ระบบโดยใช้ค่าพารามิเตอร์ กับระบบที่มี

มีการใช้อัลกอริทึม AES 128

จากภาพที่ 8 จะพบว่า การเข้าสู่ระบบโดยใช้ค่าพารามิเตอร์ โดยผ่านระบบที่ทำการใช้อัลกอริทึม AES 128 แล้ว การเข้าสู่ระบบจะล้มเหลวไม่สามารถเข้าสู่ระบบได้ แต่สำหรับผู้ใช้ที่มีสิทธิ์ใช้งานระบบและมี คีย์ ตัวเดียวกันกับทาง เซิร์ฟเวอร์ ก็จะสามารถเข้าสู่ระบบได้

3.4 การถอดรหัสข้อมูล

เมื่อได้ cipher text จากทาง client เรียบร้อยแล้ว ทางฝั่งของเซิร์ฟเวอร์ จะทำการถอดรหัสโดยใช้ คีย์ ที่ถืออยู่เหมือนกับทางฝั่งของ client หาก คีย์ ตรงกัน ผู้ใช้ก็จะสามารถใช้งานระบบได้ ซึ่งการถอดรหัสนี้ก็จะมี การถอดรหัสกลับ 2 ชั้นเช่นเดียวกับการเข้ารหัส เมื่อทางเซิร์ฟเวอร์ ทำการถอดรหัสโดยใช้ คีย์ เรียบร้อย

ภาพที่ 9: การเข้ารหัสระบบ โดยใช้การเข้ารหัสแบบ AES 128

5. สรุป

งานวิจัยนี้ได้เสนออัลกอริทึมการเข้ารหัสข้อมูล เพื่อรักษาความปลอดภัยบนระบบ เว็บแอปพลิเคชัน ภายในขององค์กร จากการโจมตีแบบ SQL Injection ผลจากการทดลองพบว่า การเข้ารหัสข้อมูลมีความปลอดภัยเพิ่มมากขึ้น โดยมีการทำงานสองขั้นตอนทำให้มีการเข้ารหัสที่ซับซ้อนเพิ่มมากขึ้นยากต่อการบุกรุก จึงส่งผลให้ผู้โจมตีไม่สามารถบุกรุกหรือดักจับข้อมูลที่เป็นความลับในระบบได้ และส่งผลให้ช่วยลดค่าใช้จ่ายขององค์กรได้ในระดับหนึ่งเนื่องจากการใช้อัลกอริทึมแทนการใช้ซอฟต์แวร์ในการป้องกันการโจมตี SQL Injection

เอกสารอ้างอิง

- [1] Web Application Security ตอนที่ 1 - 10 อันดับเทคนิคที่นิยมใช้แฮ็กเว็บไซต์ในปี 2013 <http://www.worldphp.com/blogs/518343be8e617/Web-Application-Security2013>
- [2] M. Muthuprasanna, K. Wei, and S. Kothari, "Eliminating SQL Injection Attacks - A Transparent Defense Mechanism," in *2013 15th IEEE International Symposium on Web Systems Evolution (WSE), Los Alamitos, CA, USA, 2006*, vol. 0, pp. 22–32.
- [3] S. Avireddy, V. Perumal, N. Gowraj, R. S. Kannan, P. Thinakaran, S. Ganapathi, J. R. Gunasekaran, and S. Prabhu, "Random4: An Application Specific Randomized Encryption Algorithm to Prevent SQL Injection," in *2013 12th IEEE International Conference on Trust, Security and Privacy in Computing and Communications, Los Alamitos, CA, USA, 2012*, vol. 0, pp. 1327–1333.
- [4] N. Lambert and K. S. Lin, "Use of query tokenization to detect and prevent SQL injection attacks," in *2010 3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT), 2010*, vol. 2, pp. 438–440.
- [5] D. A. Kindy and A.-S. K. Pathan, "A survey on SQL injection: Vulnerabilities, attacks, and prevention techniques," in *2011 IEEE 15th International Symposium on Consumer Electronics (ISCE), 2011*, pp. 468–471.
- [6] A. Tajpour and M. JorJor Zade Shooshtari, "Evaluation of SQL Injection Detection and Prevention Techniques," in *2010 Second International Conference on Computational Intelligencetece, Communication Systems and Networks (CICSyN), 2010*, pp. 216–221.
- [7] R. Johari and P. Sharma, "A Survey on Web Application Vulnerabilities (SQLIA, XSS) Exploitation and Security Engine for SQL Injection," in *2012 International Conference on Communication Systems and Network Technologies (CSNT), 2012*, pp. 453–458.
- [8] S. Nanda, L.-C. Lam, and T. Chiueh, "Web Application Attack Prevention for Tiered Internet Services," in *Fourth International Conference on Information Assurance and Security, 2008. ISIAS '08, 2008*, pp. 186–191.
- [9] N. Floissac and Y. L'Hyver, "From AES-128 to AES-192 and AES-256, How to Adapt Differential Fault Analysis Attacks on Key Expansion," in *2011 Workshop on Fault Diagnosis and Tolerance in Cryptography (FDTC), 2011*, pp. 43–53.
- [10] Tajpour, A., M. Masrom, M.Z. Heydari, and S. Ibrahim. "SQL Injection Detection and Prevention Tools Assessment." In *2010 3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT), 9:518522, 2010*. doi:10.1109/ICCSIT.2010.5563777.
- [11] กิตติชน แม้นสมุทร, และสุสติ บุญรอด. ระบบวิเคราะห์การโจมตีแบบเอสคิวแอล อินเจกชัน ด้วยเทคนิคพื้นฐานทางไวยากรณ์. *การประชุมทางวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ, 2550*
- [12] D. Kar and S. Panigrahi, "Prevention of SQL Injection attack using query transformation and hashing," in *Advance Computing Conference (IACC), 2013 IEEE 3rd International, 2013*, pp. 1317–1323.
- [13] สืบพงศ์ ฉัยนุ้ย. ระบบป้องกันภัยสำหรับเว็บแอปพลิเคชัน. มหาวิทยาลัยเทคโนโลยีมหานคร, 2554
- [14] Xi-Rong Wu and P. P. . Chan, "SQL injection attacks detection in adversarial environments by k-centers," 2012, pp. 406–410.
- [15] J.-G. Kim, "Injection Attack Detection Using the Removal of SQL query Attribute Values," in *2011 International Conference on Information Science and Applications (ICISA), 2011*, pp. 1–7.

การศึกษาความสัมพันธ์ของแบนด์วิดท์กับจำนวนโหนดโจมตีแบบ DDoS ด้วย NS2

A Study on the Relationship Between the Bandwidth and the Number of DDoS Attack Nodes with NS2

ประดิษฐ์ พิทักษ์เสถียรกุล (Pradit Pitaksathienkul)¹ และ ศิรปัฐ บัญครอง (Sirapat Boonkrong)²

^{1,2}ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
praditp@kmutnb.ac.th)¹, sirapatb@kmutnb.ac.th)²

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาผลกระทบของการโจมตีแบบ DDoS แบบ Flooding ด้วย NS2 โดยสร้างโมเดลการทดลองที่ทำการโจมตี โดยกำหนดค่า Throughput เท่ากัน จากนั้นปรับจำนวนโหนดที่โจมตี เพื่อพิจารณาค่าแบนด์วิดท์จากการดำเนินการวิจัย ผู้วิจัยสามารถสร้างสมการความสัมพันธ์ระหว่างแบนด์วิดท์ของผู้ถูกโจมตีและจำนวนผู้โจมตีได้ ซึ่งจะสามารถนำไปใช้ในการประเมินผลกระทบจากการโจมตีลักษณะนี้ได้

คำสำคัญ: DDoS, การโจมตี

Abstract

The objective of this research is to study the effects of the DDoS attack. The study was carried out and simulated in a network simulator, NS2. The number of attackers was adjusted so that the impact on a victim could be analysed. We obtained an equation that can help estimate the effect of such attack.

Keyword: DDoS, attacking.

1. บทนำ

1.1 สถานการณ์ปัจจุบัน

ปัจจุบัน ปัญหาการโจมตีระบบเครือข่าย เป็นปัญหาที่มีความสำคัญมาก เนื่องจากระบบเครือข่ายมีช่องทางในการสื่อสารมากขึ้นเท่าใด ผู้บุกรุกก็อาศัยช่องทางเหล่านั้น มาใช้ในการโจมตีระบบได้มากขึ้น ทำให้ภัยคุกคาม และการบุกรุกผ่านระบบเครือข่าย มีแนวโน้มเพิ่มมากขึ้น

การโจมตีแบบ DDoS (Distributed Denial of Service) เป็นการโจมตีที่สร้างปัญหาต่อประสิทธิภาพการใช้งานในระบบเครือข่ายที่มีเป้าหมายเพื่อให้ผู้ใช้ไม่สามารถใช้งานระบบเครือข่ายได้ตามปกติ หรือ สร้างความเสียหายให้กับระบบคอมพิวเตอร์ โดยอาศัยความจำกัดของทรัพยากร อย่างเช่น แบนด์วิดท์ (Bandwidth) หน่วยความจำ หรือ ความสามารถในการประมวลผล และ โดยอาศัยความเป็นระบบเปิดของระบบเครือข่าย

1.2 วัตถุประสงค์

งานวิจัยนี้ เป็นการศึกษาผลกระทบของการโจมตีแบบ DDoS แบบ Flooding โดยได้ศึกษาค่าแบนด์วิดท์ กับจำนวนโหนดที่โจมตี โดยสร้างโมเดลการทดลอง จากการปรับจำนวนโหนดที่โจมตีที่ส่งผลต่อค่าประสิทธิภาพ เพื่อสร้างสมการความสัมพันธ์ระหว่างแบนด์วิดท์กับจำนวนโหนดที่โจมตี

2. งานวิจัยที่เกี่ยวข้อง

ตามการจำแนกประเภทของการโจมตีแบบ DDoS ของ องค์กร CERT/CC แบ่งได้ 3 ประเภท [1] คือ การโจมตีแบบส่งข้อมูลจำนวนมาก การโจมตีแบบอาศัยจุดอ่อนของโปรโตคอล และ การโจมตีแบบสะท้อน [2] สำหรับงานวิจัยนี้ พิจารณา การโจมตีแบบส่งข้อมูลจำนวนมาก ซึ่งเป็นรูปแบบการโจมตีที่องค์กร CERT/CC มักได้รับการแจ้ง [3]

การโจมตีแบบส่งข้อมูลจำนวนมาก ไม่ได้มุ่งหวังให้คอมพิวเตอร์ปลายทางหยุดทำงาน แต่เป็นการขัดขวางการใช้งานระบบเครือข่ายของผู้ใช้ ทำให้ใช้งานได้ช้าลง หรือ ไม่สามารถใช้งานได้

รูปแบบการศึกษาการตรวจสอบการโจมตีแบบ DDoS มีหลายงานวิจัย แต่ละงานวิจัยต่างก็กำหนดวิธีการในการตรวจสอบการ

โจมตีแบบ DDoS แตกต่างกันไป ตามวัตถุประสงค์ที่ต้องการจะวัดค่า เช่น การตรวจสอบการโจมตี ด้วยวิธีการวัดค่าเอนโทรปี ปริมาณของข้อมูลจราจรที่มีการเปลี่ยนแปลงในช่วงเวลา [4, 5] การตรวจสอบการโจมตีด้วยวิธีนิวรัลเน็ตเวิร์ค [6, 7] การตรวจสอบการโจมตีด้วยวิธีหาค่าสหสัมพันธ์ [8] แม้วิธีการจะแตกต่างกัน แต่ผลงานวิจัยมุ่งเน้นด้านความแม่นยำในการตรวจสอบ โดยงานวิจัยนี้จัดเป็นประเภทวัดปริมาณข้อมูลจราจรในเครือข่าย ในช่วงเวลาหนึ่ง

3. วิธีการวิจัย

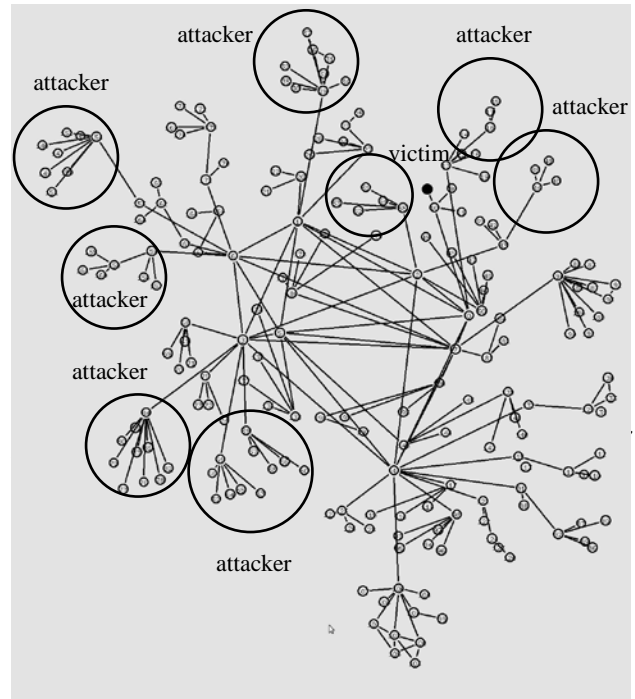
ในงานวิจัยนี้ได้สร้างแบบจำลองเครือข่าย ด้วยโปรแกรม NS2 version 2.35 กำหนดจำนวนโหนดทั้งหมด 200 โหนด ซึ่งเป็นโหนดปลายทาง รวมโหนดที่ทำหน้าที่เป็น Forwarding เพื่อให้ได้จำนวนโหนดที่รับส่งข้อมูลได้ไม่ต่ำกว่า 100 โหนด ด้วยวิธีการสุ่มใช้วิธีจำลองเครือข่ายแบบ transit-stub [10] ด้วย Waxman algorithm [11] แต่ละ link มีการรับส่งข้อมูลแบบ full duplex รองรับแบนด์วิดท์ที่ 1 Mbit/s มีค่า delay ระหว่าง 50 ถึง 500 ms มีการกำหนด queue ด้วยวิธีการ droptail

โดยโปรแกรมทำงานบนคอมพิวเตอร์ CPU Intel Core2 Duo 2.28 GHz หน่วยความจำ 2GB ระบบปฏิบัติการ Ubuntu version 12.04

3.1 แบบจำลองการทดลอง

กำหนดเครื่องที่เป็นเป้าหมายของการโจมตี (victim สัญลักษณ์เป็นจุดสีดำในภาพ) ให้มีการทำงานส่งข้อมูลตามปกติ 1 โหนด ให้มีการส่งข้อมูลแบบ udp และ tcp จากโหนดที่อยู่ข้างเคียง ไปยังเครื่องที่เป็นเป้าหมายของการโจมตี ขนาดของ packet เป็น 100 bytes ด้วยอัตราการส่งถ่ายข้อมูล 100 K packets ต่อวินาที ตามลำดับ เริ่มจากวินาทีที่ 10 ถึงวินาทีที่ 80 ทำการจับเวลา ณ โหนดที่รับข้อมูล ทุก 0.5 วินาที เพื่อคำนวณแบนด์วิดท์

$\text{Bandwidth} = \text{ปริมาณข้อมูลที่รับได้} / 0.5 * 8 / 1,000,000$
หน่วยที่ได้ เป็น Mbit/วินาที



ภาพที่ 1: แบบจำลองโหนดปกติ และ โหนดโจมตี

3.2 วิธีการทดลอง

ที่เวลา 20 วินาที กำหนดให้เริ่มมีการโจมตี จากกลุ่มของโหนดที่วงกลมไว้ โดยแบ่งจำนวนโหนดการโจมตี เพิ่มขึ้นครั้งละ 10 โหนด จนถึง 50 โหนด ซึ่งทำให้เครื่องเป้าหมายจะไม่สามารถส่งข้อมูลปกติไปยังโหนดข้างเคียงได้ (เข้าใกล้ 0)

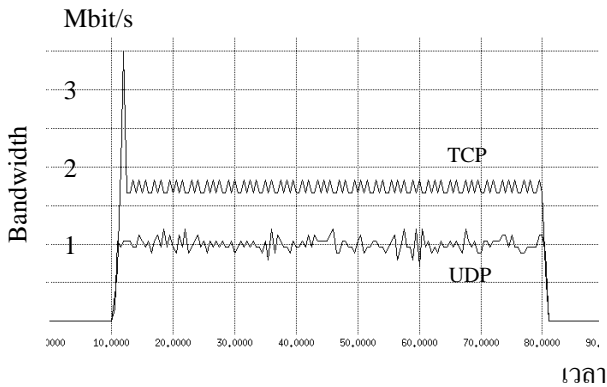
ในการโจมตีนั้น กำหนดให้แต่ละโหนดโจมตี มีการส่งข้อมูลชนิด udp ด้วยขนาดของ packet เป็น 100 bytes ด้วยอัตราการส่งถ่ายข้อมูล 40 K packets ต่อวินาที ตามลำดับ

ทำการคำนวณค่าแบนด์วิดท์ ทั้งที่เป็นแบนด์วิดท์ของการส่งข้อมูลปกติ และ แบนด์วิดท์ของการส่งข้อมูลโจมตี ที่เครื่องเป้าหมายเพื่อหาค่าความสัมพันธ์

4. ผลการทดลอง

4.1 ผลการทดลองเบื้องต้น

เริ่มจากขณะที่ยังไม่มีการโจมตี ความสัมพันธ์ของแบนด์วิดท์ กับเวลา ในการส่งข้อมูล เป็นดังภาพที่ 2

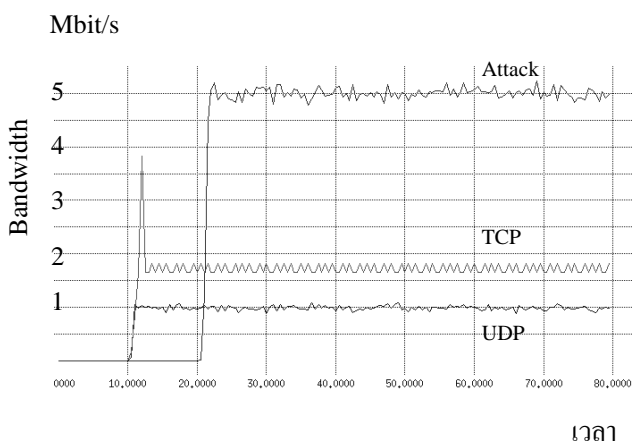


ภาพที่ 2: กราฟความสัมพันธ์ของ แบนด์วิดท์ กับ เวลาของการรับส่งข้อมูลแบบ UDP และ TCP ขณะที่ยังไม่มีการโจมตี

เส้นกราฟ UDP หมายถึง อัตราการรับส่งข้อมูลเฉลี่ย แบบ UDP ณ เครื่องเป้าหมาย วัดได้ เท่ากับ 1 Mbit/s

เส้นกราฟ TCP หมายถึง อัตราการรับส่งข้อมูลเฉลี่ย แบบ TCP ณ เครื่องเป้าหมาย วัดได้ เท่ากับ 1.8 Mbit/s ซึ่งต่อไปจะแสดงเส้นกราฟ TCP เพื่อให้เห็นความเปลี่ยนแปลงที่เกิดขึ้นแต่ไม่ได้นำมาใช้ในการคำนวณ

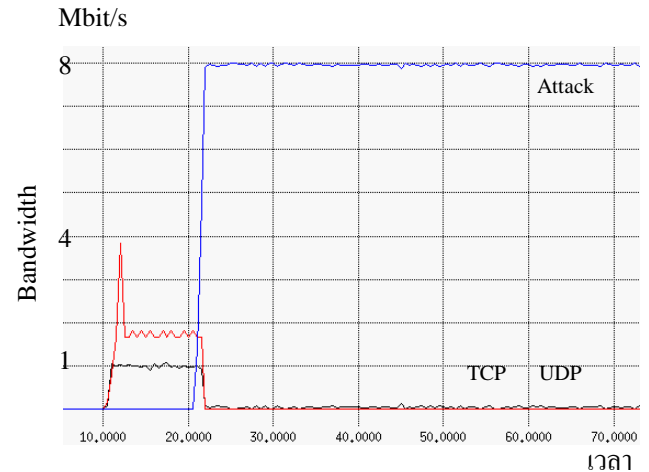
เมื่อเวลา 20 วินาที กำหนดให้ โหนดโจมตีจำนวน 10 โหนด ส่งข้อมูลโจมตีแบบ udp มายังเครื่องเป้าหมาย ผลที่เกิดขึ้น เป็นตามภาพที่ 3



ภาพที่ 3: กราฟความสัมพันธ์ของ แบนด์วิดท์ กับ เวลาขณะที่มีการโจมตี จำนวน 10 โหนด

เส้นกราฟ UDP หมายถึง อัตราการรับส่งข้อมูลเฉลี่ย แบบ UDP ณ เครื่องเป้าหมาย วัดได้ เท่ากับ 0.98 Mbit/s

เส้นกราฟ Attack หมายถึง อัตราการรับส่งข้อมูลเฉลี่ย แบบ UDP ณ เครื่องเป้าหมาย วัดได้ เท่ากับ 4.91 Mbit/s
ทำการเพิ่มจำนวน โหนดโจมตีขึ้นเรื่อยๆ จนถึง 50 โหนดโจมตี



ภาพที่ 4: กราฟความสัมพันธ์ของ แบนด์วิดท์ กับ เวลาขณะที่มีการโจมตี จำนวน 50 โหนด

เส้นกราฟ UDP หมายถึง อัตราการรับส่งข้อมูลเฉลี่ย แบบ UDP ณ เครื่องเป้าหมาย วัดได้ เท่ากับ 0.19 Mbit/s

เส้นกราฟ Attack หมายถึง อัตราการรับส่งข้อมูลเฉลี่ย แบบ UDP ณ เครื่องเป้าหมาย วัดได้ เท่ากับ 7.81 Mbit/s

จากการทดลองพบว่า เมื่อเพิ่มจำนวน โหนดโจมตี จะทำให้ แบนด์วิดท์ที่ใช้งานปกติ ลดลง แสดงผลได้ดัง ตารางที่ 1

ตารางที่ 1: แสดงความสัมพันธ์ของจำนวน โหนดที่โจมตี กับ แบนด์วิดท์ที่ใช้งานปกติ

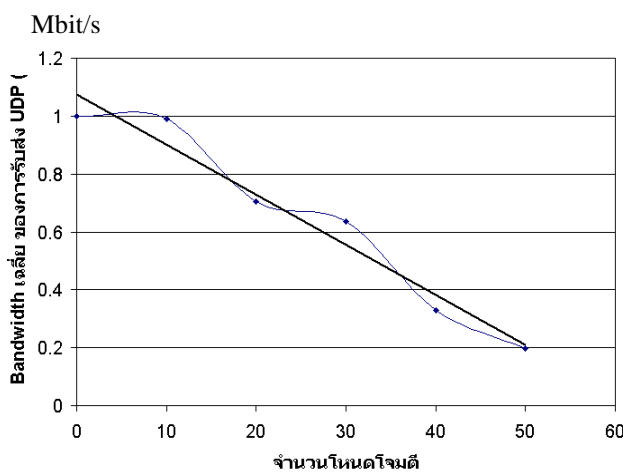
จำนวน โหนดโจมตี	แบนด์วิดท์ของการส่งข้อมูลแบบ UDP ที่ใช้งานปกติ (Mbit/s)	แบนด์วิดท์ลดลงจากขณะที่ไม่มีการโจมตีคิดเป็นเปอร์เซ็นต์ (%)
0	1	0
10	0.98	-1.13
20	0.70	-29.41
30	0.63	-36.54
40	0.32	-67.05
50	0.19	-80.36

จากตารางที่ 1 เมื่อคำนวณแบนด์วิดท์ที่ใช้งานปกติ ลดลง เป็นเปอร์เซ็นต์ ได้ค่าดังนี้

- เมื่อเพิ่มจำนวนโหนดโจมตี เป็น 10 โหนด แบนด์วิดท์ที่ใช้งานปกติ ลดลงจากขณะที่ไม่มีการโจมตี เท่ากับ -1.13%
- เมื่อเพิ่มจำนวนโหนดโจมตี เป็น 20 โหนด แบนด์วิดท์ที่ใช้งานปกติ ลดลงจากขณะที่ไม่มีการโจมตี เท่ากับ -29.41%
- เมื่อเพิ่มจำนวนโหนดโจมตี เป็น 30 โหนด แบนด์วิดท์ที่ใช้งานปกติ ลดลงจากขณะที่ไม่มีการโจมตี เท่ากับ -36.54%
- เมื่อเพิ่มจำนวนโหนดโจมตี เป็น 40 โหนด แบนด์วิดท์ที่ใช้งานปกติ ลดลงจากขณะที่ไม่มีการโจมตี เท่ากับ -67.05%
- เมื่อเพิ่มจำนวนโหนดโจมตี เป็น 50 โหนด แบนด์วิดท์ที่ใช้งานปกติ ลดลงจากขณะที่ไม่มีการโจมตี เท่ากับ -80.36%

4.2 ความสัมพันธ์ระหว่างแบนด์วิดท์ กับ จำนวนโหนดโจมตี

จากผลการทดลองเบื้องต้น สามารถนำแบนด์วิดท์เฉลี่ยในช่วงระยะเวลาทำการทดลองที่ได้จากแต่ละการทดลอง มาสร้างกราฟของความสัมพันธ์ ได้ดังนี้



ภาพที่ 5: กราฟแสดงความสัมพันธ์ระหว่างแบนด์วิดท์ กับ จำนวนโหนดโจมตี

จากกราฟ เส้นโค้งที่เกิดขึ้น เป็นผลจากการ plot ค่าแบนด์วิดท์ กับจำนวนโหนดโจมตี

ส่วนเส้นตรง เป็นเส้นแนวโน้มที่คำนวณขึ้น

เมื่อพิจารณาความชันของเส้นแนวโน้ม แสดงให้เห็นถึงความสามารถในการส่งข้อมูลใช้งานตามปกติ จะลดลง เมื่อจำนวนโหนดโจมตี เพิ่มขึ้น

จากภาพที่ 5 สามารถกำหนดสมการความสัมพันธ์ระหว่างแบนด์วิดท์ของโหนดที่ถูกโจมตี และจำนวนผู้โจมตี ได้ดังนี้

$$y = -0.017x + 1.075$$

โดยกำหนดให้ y คือ แบนด์วิดท์ที่เป็นการส่งข้อมูล ตามปกติ และ x คือ จำนวนโหนดโจมตี

จากสมการที่ได้ดังกล่าว สามารถสร้างความสัมพันธ์ของแบนด์วิดท์ กับ จำนวนโหนดโจมตีได้ โดยถ้าทราบจำนวนโหนดโจมตี จะสามารถคำนวณค่าแบนด์วิดท์ที่เป็นการส่งข้อมูลตามปกติ ลดลงได้ ภายใต้สภาพแวดล้อมที่ได้กำหนดขึ้นในการทดลอง

สมการที่ได้นี้ ผู้วิจัยเห็นว่า จะมีประโยชน์ในการประเมินผลกระทบที่อาจจะเกิดขึ้น ได้บนระบบเครือข่าย หรือต่อโหนดบนระบบเครือข่ายนั้น ๆ ต่อจำนวนผู้โจมตี ซึ่งในอนาคตจะสามารถพัฒนาขึ้นให้เป็น Dynamic มากขึ้น เพื่อที่ผู้ออกแบบระบบเครือข่ายจะสามารถนำไปประมาณการ หรือประเมินผลกระทบล่วงหน้าได้

5. สรุป

งานวิจัยนี้ มีวัตถุประสงค์เพื่อศึกษาผลกระทบของการโจมตีแบบ DDoS แบบ Flooding โดยศึกษาค่าแบนด์วิดท์ที่ใช้งานปกติ กับจำนวนโหนดที่โจมตี จากการสร้างโมเดลการทดลอง จากการปรับจำนวนโหนดที่โจมตีที่ส่งผลต่อค่าประสิทธิภาพ เพื่อสร้างสมการความสัมพันธ์ระหว่างแบนด์วิดท์กับจำนวนโหนดที่โจมตี นั้น

ผลการทดลองชี้ให้เห็นได้ว่า การโจมตีแบบส่งข้อมูลจำนวนมากที่ใช้ UDP มีผลกระทบต่อการใช้งานระบบเครือข่าย ที่เป็นการรับส่งข้อมูลอย่างเห็นได้ชัด เมื่อพิจารณาแบนด์วิดท์ ในขณะที่การรับส่งข้อมูลแบบ UDP เป็นการทำงานแบบ Connectionless-Oriented ซึ่งไม่ต้องการการตอบกลับจากเครื่องปลายทาง หรือ เป็นการส่งทางเดียว ทำให้สามารถส่งข้อมูลไปได้ ตลอดเวลาอย่างต่อเนื่อง สามารถนำมาใช้วัดประสิทธิภาพ ของการรับส่งข้อมูลในระบบเครือข่ายได้

จากผลการทดลอง สามารถสรุปได้ว่า จำนวนโหนดที่ใช้ในการโจมตี มีผลต่อแบนด์วิดท์ที่ใช้งานปกติ เป็นแบบปฏิภาคผกผันกัน

เอกสารอ้างอิง

- [1] N. Bais, N. Titarmare, A. B. Patil, "A Survey : Detection and Defense Mechanism against DDoS Attacks", *IEEE Communications Surveys & Tutorials*, vol. 15, no. 4, pp. 2046-2069, 2013.
- [2] J. Mirkovic, "A Taxonomy of DDoS Attack and DDoS Defense Mechanisms", *ACM SIGCOMM Computer Communications Review*, vol. 34, no. 2, pp. 39-53, 2004.
- [3] K. J. Houle, G. M. Weaver, "Trends in Denial of Service Attack Technology", *CERT Coordination Center*, 2001.
- [4] J. H. Jun, H. Oh, "DDoS flooding attack detection through a step-by-step investigation", *IEEE Networked Embedded Systems for Enterprise Applications (NESEA)*, pp.1-5, 2011.
- [5] Y. Tao, S. Yu, "DDoS Attack Detection at Local Area Networks Using Information Theoretical Metrics", *IEEE Trust, Security and Privacy in Computing and Communications (TrustCom)*, pp. 233-240, 2013.
- [6] J. Li, Y. Liu, L. Gu, "DDoS attack detection based on neural network", *Aware Computing (ISAC)*, pp. 196-199, 2010.
- [7] A. Sanmorino, S. Yazid, "DDoS Attack Detection Method and Mitigation Using Pattern of the Flow", *Information and Communication Technology (ICoICT)*, pp. 12-16, 2013.
- [8] L. Rua, R. Guang-min, Y. Xing-miao, "Global Detection of DDoS Attack Based on Time and Frequency Analysis", *ICCCAS Communications, Circuits and Systems*, pp. 462-466, 2007.
- [9] R. Braga, E. Mota, A. Passito, "Lightweight DDoS flooding attack detection using NOX/OpenFlow", *IEEE Conference on Local Computer Networks*, 2010.
- [10] E. W. Zegura, K. L. Calvert, S. B. acharjee, "How to model an internet network", *INFOCOM '96. Fifteenth Annual Joint Conference of the IEEE Computer Societies*, vol.2, pp. 594-602, 1996.
- [11] B.M. Waxman, "Routing of multipoint connections", *IEEE Journal on Selected Areas in Communications*, vol. 6, no. 9, pp.1617-1622, 1988.
- [12] จิตติมา นิตยวรรณ, โสภส ชกัตตราชางษ์, อุไรรัตน์ พึ่งสุนทรบัตร์, "Network Simulator (NS2)", *ศูนย์เทคโนโลยีอิเล็กทรอนิกส์ และคอมพิวเตอร์แห่งชาติ*.
- [13] ศุภโชค สุขเกษม, "การวิเคราะห์ข้อมูลกิจกรรมของระบบเพื่อตรวจจับการบุกรุก", *มหาวิทยาลัยสงขลานครินทร์*, 2548.
- [14] เรืองไกร รังสิพล, "เจาะระบบ TCP/IP จุดอ่อนของโปรโตคอลและวิธีป้องกัน", *สำนักพิมพ์โปรวิชั่น*, 2544.
- [15] DDoS Top 2011, <http://www.net-security.org>, 2011.

การจัดเก็บรหัสผ่านอย่างปลอดภัยโดยใช้เทคนิคการปรับค่าซอลต์ที่เหมาะสม ร่วมกับแฮชฟังก์ชัน

Secure Password Storing using Dynamic Salt Selection with Hash Function

ชาวลิต สมบูรณ์พัฒนากิจ (Chaowalit Somboonpattanakit)¹ และศิริปัฐ บุญครอง (Sirapat Boonkrong)²

^{1,2}ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹admin@rmutp.ac.th, ²sirapatb@kmutnb.ac.th

บทคัดย่อ

การใช้งาน *Cryptographic hashes* ถือเป็นหนึ่งในวิธีการที่ได้รับความนิยมในการจัดการด้านการรักษาความปลอดภัยของแอปพลิเคชัน เช่น กระบวนการ *Authentication*, การสร้าง *Digital signature*, การทำ *Message integrity codes* รวมถึงการจัดเก็บรหัสผ่าน ซึ่ง *Hash Function* นั้นมีความเร็วในการทำงานสูง จึงทำให้สามารถถูกโจมตีด้วยการใช้ *rainbow table* หรือ *dictionary attack* ได้ ทำให้วิธีการในการจัดเก็บรหัสผ่านโดยการนำรหัสผ่านมาผ่านกระบวนการของ *Hash function* เพียงอย่างเดียวไม่สามารถที่จะรักษาความปลอดภัยของรหัสผ่านได้อีกต่อไป งานวิจัยนี้จึงทำการพัฒนาอัลกอริทึมที่สามารถเลือกใช้ค่าซอลต์ที่เหมาะสมและกำหนดค่ารูปแบบเพื่อใช้ในการแทรกซอลต์ลงในรหัสผ่าน เพื่อให้สามารถรักษาความปลอดภัยของรหัสผ่านที่จัดเก็บลงในฐานข้อมูลและมีความทนทานต่อการโจมตี

คำสำคัญ: รหัสผ่าน แฮชฟังก์ชัน ซอลต์ ความปลอดภัย
เรนโบว์เทเบิล

Abstract

Cryptographic hashes are one of the most popular methods used in Security applications. They are used in authentication, digital signatures, message integrity codes and password storing. Hash functions are designed to be very fast, which makes them vulnerable to rainbow table and dictionary attacks. This means that storing passwords using only hash function is not enough to prevent the attack. This research, therefore, propose a new algorithm That helps select a suitable salt value for each password in orders to increase its security.

Keyword: password, hash function, salt, secure, rainbow table.

1. บทนำ

ปัจจุบันการใช้งานรหัสผ่านถือเป็นมาตรการหนึ่งของความปลอดภัยขั้นพื้นฐานที่นิยมใช้กันอย่างแพร่หลายเพื่อใช้ในการตรวจสอบสิทธิ์หรือการยืนยันตัวตนเพื่อเข้าใช้งานระบบต่างๆ [1], [2] อย่างไรก็ตาม รหัสผ่านที่จัดเก็บไว้อย่างเป็นความลับนั้น อาจจะไม่เป็นความลับหารหัสผ่านดังกล่าวถูกผู้อื่นล่วงรู้และนำไปใช้ในทางมิชอบ ในการกำหนดรหัสผ่านยังมีกลไกที่สามารถนำมาควบคุมและสร้างข้อจำกัดเพื่อความปลอดภัยยิ่งขึ้น เช่น การกำหนดอายุการใช้งานของรหัสผ่าน การบังคับให้ตั้งรหัสผ่านใหม่เมื่อครบระยะเวลา การกำหนดให้ตั้งรหัสผ่านที่ยากต่อการคาดเดา [3] แต่ก็มีผู้ใช้งานไม่น้อยเลือกที่จะตั้งรหัสผ่านชุดเดียวกันสำหรับทุกระบบที่ตนสามารถเข้าใช้งาน เนื่องจากต้องการความสะดวกและง่ายต่อการจดจำ เพราะฉะนั้นผู้ที่ให้บริการหรือผู้พัฒนาระบบหากไม่คำนึงถึงเรื่องของการรักษาข้อมูลความลับของผู้ใช้งาน หากมีการโจรกรรมข้อมูลเกิดขึ้น นั่นหมายถึงความเสียหายที่จะเกิดขึ้นต่อผู้ใช้นั้น ไม่ได้เสียหายต่อระบบที่ดูแลรับผิดชอบเพียงระบบเดียว แต่อาจจะหมายถึงระบบอื่นๆที่ผู้ใช้มีสิทธิ์ในการเข้าใช้งานอีกด้วย [4] ความมั่นคงปลอดภัยข้อมูลจึงถือเป็นเรื่องสำคัญอย่างมาก เมื่อเกิดข้อผิดพลาดอาจนำมาซึ่งความเสียหายที่ไม่อาจประเมินค่าได้ ดังนั้น ถ้าภายในระบบมีการควบคุมความปลอดภัยที่ดี จะช่วยลดโอกาสเสี่ยงหรือความน่าจะเป็นที่จะเกิดเหตุการณ์การที่ถูกคุกคามหรือโจรกรรมผ่านระบบได้

งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อสร้างอัลกอริทึมในการเลือกใช้ค่าซอลต์ที่เหมาะสมและสามารถกำหนดค่ารูปแบบที่

ใช้ในการแทรกซอลท์ลงในรหัสผ่าน เพื่อให้สามารถรักษาความลับของรหัสผ่านที่จัดเก็บลงในฐานข้อมูลให้มีความทนทานต่อการโจมตี ซึ่งงานวิจัยนี้จัดลำดับหัวข้อตามลำดับดังนี้ คือ ส่วนที่ 2 อธิบายถึงทฤษฎีและงานวิจัยที่เกี่ยวข้อง ส่วนที่ 3 วิธีดำเนินการวิจัย ส่วนที่ 4 ผลการดำเนินงาน และบทสรุปของงานวิจัยนี้อยู่ในส่วนที่ 5

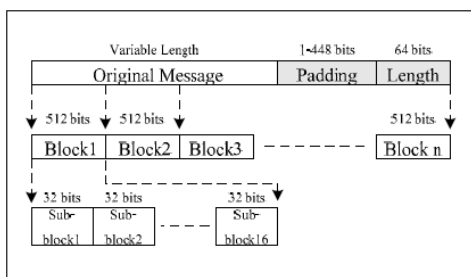
2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 แฮชฟังก์ชัน

คือกระบวนการทางด้านคณิตศาสตร์ที่ใช้ในการย่อข้อมูล (message digest) ไม่ว่าจะเป็นไฟล์เอกสาร, Plaintext หรืออื่นๆ โดยผลลัพธ์ที่ได้จากการย่อนี้จะถูกเรียกว่า ค่าแฮช (hash value) [5] ซึ่งค่าแฮชนี้มีความยาวแตกต่างกันไปขึ้นอยู่กับอัลกอริทึมที่ใช้ เช่นค่าแฮชขนาด 128 บิตสำหรับอัลกอริทึม MD5 และ 160 บิตสำหรับอัลกอริทึม SHA1 เป็นต้น

2.2 MD5 อัลกอริทึม

การทำงานของอัลกอริทึม MD5 จะมีการแบ่งข้อความต้นฉบับขนาดใด ๆ ออกเป็นกลุ่มบิตหลาย ๆ กลุ่มบิต ที่มีขนาด 512 บิต ตามลำดับ แต่ละกลุ่มบิตนี้จะถูกแบ่งย่อยออกเป็น 16 กลุ่มย่อย (Sub-block) กลุ่มบิตย่อยละ 32 บิต และเมื่อผ่านการดำเนินการตามที่กำหนดไว้ในอัลกอริทึม MD5 จนครบแล้วจะให้ผลลัพธ์เป็นเซตของกลุ่มบิตย่อยขนาด 32 บิต จำนวน 4 กลุ่ม เพื่อนำมารวมกันเป็นค่าแฮชผลลัพธ์ ขนาด 128 บิต [6] ดังแสดงในภาพที่ 1 ซึ่ง MD5 เป็นอัลกอริทึมที่นิยมนำมาใช้ในการจัดเก็บรหัสผ่าน



ภาพที่ 1 : แสดงการจัดเตรียมชุดข้อมูลสำหรับอัลกอริทึม MD5

2.3 การจัดเก็บรหัสผ่าน

เทคนิคและวิธีการที่ใช้ในการจัดเก็บรหัสผ่านสามารถจัดเก็บได้หลากหลายวิธี ดังต่อไปนี้ [7], [8], [9]

1. การจัดเก็บแบบ plaintext password โดยวิธีการนี้จะไม่มี การเข้ารหัสใดๆ ซึ่งวิธีนี้เป็น การเก็บแบบที่ง่ายที่สุดแต่จะมีความปลอดภัยต่ำที่สุดเนื่องจากเมื่อผู้โจมตีสามารถเข้าถึงฐานข้อมูลที่เกี่ยวข้องได้ก็จะสามารถมองเห็นรหัสผ่านนั้นโดยทันที

2. การจัดเก็บแบบ hash password โดยวิธีการนี้จะทำการจัดเก็บรหัสผ่านด้วยการนำค่ารหัสผ่านไปผ่านกระบวนการของ hash function ก่อน เช่น MD5 หรือ SHA1 แล้วจึงเก็บค่าแฮชที่ได้ไว้ในฐานข้อมูล การเก็บรหัสผ่านด้วยวิธีนี้เมื่อผู้โจมตีสามารถเข้าถึงฐานข้อมูลได้จะมองเห็นแต่เพียงค่าแฮชเท่านั้น แต่วิธีนี้สามารถถูกโจมตีได้โดยการใช้เรนโบว์เทเบิลเพื่อถอดค่าแฮชให้กลับเป็นค่ารหัสผ่าน

3. การจัดเก็บแบบ Double hash password โดยวิธีการนี้จะเหมือนกันกับวิธี hash password แต่จะแตกต่างกันตรงที่จะมีการนำค่าแฮชที่ได้ในรอบที่ 1 มาทำการแฮชซ้ำอีก 1 รอบ โดยในรอบที่ 2 อาจจะมีการเปลี่ยนแฮชฟังก์ชันหรือใช้แฮชฟังก์ชันเหมือนกับในรอบแรกก็ได้ ซึ่งวิธีนี้เมื่อผู้โจมตีสามารถเข้าถึงฐานข้อมูลได้จะมองเห็นแต่เพียงค่าแฮชเท่านั้น วิธีนี้สามารถป้องกันได้ดีกว่าการแฮชรอบเดียว แต่ผู้โจมตีสามารถทำการทดสอบเพื่อหาค่าของจำนวนรอบที่ต้องใช้ในการแฮชแล้วสร้างเรนโบว์เทเบิลใหม่ขึ้นมาโจมตีได้

4. การจัดเก็บแบบ Salted hash password โดยวิธีการนี้จะมีการเติมค่าซอลท์ลงไปในรหัสผ่านก่อนแล้วจึงเข้าสู่แฮชฟังก์ชัน โดยค่าซอลท์หมายถึง ค่าที่ถูกสร้างขึ้นมาโดยใช้การกำหนดค่าแบบคงที่หรือใช้วิธีการสุ่ม ทำให้ผลลัพธ์หรือค่าแฮชที่ได้มีความแตกต่างกันและทำให้การโจมตีด้วยเรนโบว์เทเบิลนั้นทำได้ยากขึ้น แต่เนื่องจากค่าซอลท์ที่สร้างขึ้นจะถูกจัดเก็บลงในฐานข้อมูลด้วยเสมอ ซึ่งในกรณีที่ผู้โจมตีทราบค่าซอลท์นั้นก็ยังสามารถทำการโจมตีได้เช่นกัน

2.4 เรนโบว์เทเบิล

เรนโบว์เทเบิลเป็นตารางขนาดใหญ่ที่ถูกสร้างขึ้นเพื่อจัดเก็บค่ารหัสผ่านและค่าแฮชของรหัสผ่านนั้นเพื่อนำมาใช้ในการเปรียบเทียบเวลาที่นำไปใช้ในการโจมตีเนื่องจากแฮชฟังก์ชันส่วนมากจะเป็น one-way hash function ซึ่งไม่สามารถทำกระบวนการย้อนกลับเพื่อหาผลลัพธ์ได้ [1] และการโจมตีด้วยเทคนิคนี้จะใช้เวลาน้อยกว่าเมื่อเปรียบเทียบกับวิธี brute force

แม้ว่าการสร้างเรนโบว์เทเบิลนั้นจะใช้พื้นที่ในการจัดเก็บค่อนข้างมาก [10] แต่เรนโบว์เทเบิลที่สร้างไว้แล้วสามารถนำกลับมาใช้ซ้ำได้

2.5 งานวิจัยที่เกี่ยวข้อง

การสร้างรหัสผ่านด้วยวิธีการสุ่มนั้นไม่สามารถชีวิตได้ว่ารหัสผ่านที่ได้นั้นมีคุณภาพที่ดี เพราะการวัดว่ารหัสผ่านที่มีอยู่นั้นมีคุณภาพที่ดีหรือไม่นั้นควรดูจากเวลาที่ต้องใช้ในการโจมตีรหัสผ่านนั้น ซึ่งได้ระบุไว้ว่ารหัสผ่านที่ดี จะต้องเป็นรหัสผ่านที่มีคุณสมบัติดังต่อไปนี้คือ มีความยาวอย่างน้อย 8 ตัวอักษรและประกอบไปด้วยอักขระพิเศษ 3 ตัวอักษร และในรหัสผ่านควรจะเป็นตัวอักษรผสมกับตัวเลข [11] โดยในปี 2012 มีนักวิจัยชาวจีน ได้ให้ความสนใจในการศึกษาเกี่ยวกับแอปพลิเคชันในปัจจุบันและการรักษาความปลอดภัยของรหัสผ่านโดยใช้ MD5 algorithm สำหรับกระบวนการทำ authentication ซึ่งมีแนวคิดในการทำให้ MD5 hash function สามารถจัดเก็บรหัสผ่านได้ปลอดภัยยิ่งขึ้น [12] โดยเทคนิคที่นำเสนอคือ การปรับเปลี่ยน MD5 processing ให้ต้องใช้เวลาในการประมวลผลที่นานขึ้น การแทรกข้อมูลเข้าไปในรหัสผ่าน โดยการต่อที่ด้านหน้าการแบ่งข้อมูลออกเป็น 2 ส่วนก่อนทำการแทรกข้อมูลลงไป ในรหัสผ่าน ซึ่งการวัดประสิทธิภาพของการเติมซอลต์ลงในรหัสผ่าน [8] โดยการทำให้เป็น prefix salt หรือ suffix salt นั้นสามารถทำให้ผู้โจมตีคาดเดาได้ยากขึ้นแต่ถ้าผู้โจมตีทำการทดลองหลายๆครั้งจนเจอ fix point แล้วก็จะไม่สามารถรักษาความปลอดภัยของรหัสผ่านได้

3. วิธีดำเนินการวิจัย

ในงานวิจัยนี้ทำการศึกษาเกี่ยวกับวิธีการจัดเก็บรหัสผ่านให้มีความปลอดภัยและทนทานต่อการโจมตี ซึ่งประกอบไปด้วย 3 ส่วนการทำงานหลักคือ การปรับคุณภาพของรหัสผ่านให้เป็นรหัสผ่านที่มีความแข็งแรง [11] การเลือกใช้ซอลต์ที่มีความเหมาะสม และการสร้างค่ารูปแบบสำหรับการแทรกซอลต์ลงในรหัสผ่าน

3.1 การปรับคุณภาพของรหัสผ่าน

ในขั้นตอนนี้จะเป็นการปรับคุณภาพของรหัสผ่านตั้งต้นให้

เป็นรหัสที่มีความแข็งแรง ซึ่งมีวิธีการทำงานดังนี้ เมื่อรับรหัสผ่านตั้งต้นเข้ามาจะทำการตรวจสอบคุณสมบัติคือ ต้องมีความยาวอย่างน้อย 8 ตัวอักษร ประกอบไปด้วยอักขระพิเศษ 3 ตัวอักษร และในรหัสผ่านต้องเป็นตัวอักษรผสมกับตัวเลข และในกรณีที่รหัสผ่านตั้งต้นมีคุณสมบัติไม่ครบก็จะทำการสร้างซอลต์ที่ขึ้นเพื่อปรับความแข็งแรงของรหัสผ่านให้เป็นไปตามเกณฑ์ดังกล่าว โดยจะใช้ชุดข้อมูลที่ประกอบด้วยตัวอักษร ตัวเลขและตัวอักขระพิเศษ (Mix alpha-numeric-all) ดังแสดงในตารางที่ 1

ตารางที่ 1: ชุดตัวอักษรที่ใช้ในการสร้างซอลต์

Dataset	Character
Special	!@#\$%^&*()-_+=~`[]{} :;'"<>.,?/
Numeric	0123456789
Alphabet	abcdefghijklmnopqrstuvwxyz ABCDEFGHIJKLMNOPQRSTUVWXYZ

3.2 การเลือกใช้ซอลต์ที่มีความเหมาะสม

เพื่อหาค่าความยาวของซอลต์ที่เหมาะสม ผู้วิจัยได้ทำการศึกษาเวลาที่ใช้ในการสร้างเรนโบว์เทเบิลเพื่อโจมตีรหัสผ่านที่มีความยาวตั้งแต่ 6 ถึง 11 ตัวอักษร และเวลาที่ใช้ในการแฮชรหัสผ่านรวมกับซอลต์ที่ความยาวตั้งแต่ 0 ถึง 128 ตัวอักษร เพื่อนำมาใช้ในการกำหนดค่าความยาวของซอลต์

1. เวลาที่ใช้ในการสร้าง เรนโบว์เทเบิล ซึ่งเครื่องมือที่ใช้ทดสอบคือเครื่องคอมพิวเตอร์ที่มีการ์ดประมวลผลกราฟิก GPU ATI Radeon 7970 จำนวน 2 การ์ด ซึ่งมีความเร็วในการคำนวณแซฟท์แวร์ที่เร็วกว่าซีพียู [13] ในการสร้างเรนโบว์เทเบิลซึ่งผลการทดสอบคือสามารถสร้างเรนโบว์เทเบิลของรหัสผ่านความยาว 6 ตัวอักษรได้ภายใน 47 วินาที และที่ความยาว 7 ตัวอักษรจะใช้เวลา 1.14 ชั่วโมง ส่วนที่ความยาว 8 ตัวอักษรใช้เวลาถึง 465 วัน และที่มีความยาวตั้งแต่ 9 ตัวอักษรขึ้นไปยังไม่สามารถคำนวณเวลาที่ใช้ในการสร้างเรนโบว์เทเบิลได้ ดังแสดงในตารางที่ 2

ตารางที่ 2: เวลาที่ใช้ในการสร้าง Rainbow Table ด้วย GPU

Password length (ตัวอักษร)	Precomputation time
1 - 6	47 seconds
1 - 7	1.14 hours
1 - 8	465 days
1 - 9	∞

2. เวลาที่ใช้ในการแฮช โดยทำการทดสอบเพื่อหาค่าเวลาที่
 ต้องใช้ในการแฮชรหัสผ่านร่วมกับซอลท์ที่ระดับความยาว
 ต่างๆ โดยใช้ข้อมูลรหัสผ่านจากเว็บไซต์ที่ได้ทำเก็บรวบรวมว่า
 เป็นรหัสผ่านที่มีผู้ใช้นิยมมากที่สุด ประกอบไปด้วยรหัสผ่าน
 ทั้งหมด 10,000 รหัสผ่าน [14] ซึ่งจากผลการทดสอบ ผู้วิจัยจึงทำ
 การสร้างค่าซอลท์เพิ่มเติมด้วยวิธีการสุ่ม โดยจะกำหนดค่าความ
 ยาวของซอลท์ตั้งแต่ 10 ถึง 32 ตัวอักษร เนื่องจากใช้เวลาในการ
 แฮชที่เหมาะสมและมีความยาวเพียงพอที่สามารถทนทานจาก
 การโจมตีด้วยเรนโบว์เทเบิล ดังแสดงในตารางที่ 3

ตารางที่ 3: เวลาที่ใช้ในการแฮชรหัสผ่านร่วมกับซอลท์ที่ระดับความยาว
 ต่างๆ

Salt (bit)	characters	Hashtime (ms)
No salt	0	0.002707
64 bit	8	0.002934
80 bit	10	0.002945
128 bit	16	0.002959
256 bit	32	0.002982
512 bit	64	0.003102
1024 bit	128	0.003279

3.3 การสร้างค่ารูปแบบสำหรับการแทรกซอลท์

การสร้างค่ารูปแบบสำหรับการแทรกซอลท์ ด้วยการใช้
 oneway-algorithm โดยมีวัตถุประสงค์คือไม่ต้องการให้สามารถ
 ทำกระบวนการย้อนกลับเพื่อคำนวณหาค่ารูปแบบนี้ได้ โดยนำ
 รหัสผ่านตั้งต้น ไปผ่านกระบวนการแฮชฟังก์ชัน ซึ่งในที่นี้จะใช้
 MD5 ซึ่งจะ ได้ผลลัพธ์เป็นค่าแฮชที่มีความยาว 128 บิต ซึ่ง
 สามารถเขียนให้อยู่ในรูปแบบของเลขฐาน 16 ได้ 32 ตัวอักษร
 เรียงต่อกัน จากนั้นทำการแปลงค่ารหัสผ่านตั้งต้นและค่าแฮช
 นั้นให้เป็นไบนารีแล้วนำบิตสุดท้ายมาทำ Exclusive OR กัน ดัง
 ตัวอย่างในตารางที่ 4

ตารางที่ 4: ผลลัพธ์การ XOR ระหว่างรหัสผ่านตั้งต้นและค่าแฮช

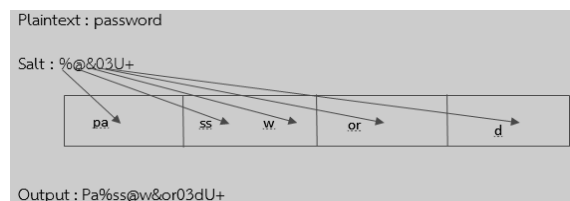
#	Binary (pass = pass) (hash = d41d)			
pass	01110000	01100001	01110011	01110011
hash	01100100	00110100	00110001	01100100
output	0	1	0	1

จากนั้นจะทำการแทรกซอลท์ลงในรหัสผ่านตามค่ารูปแบบ
 ที่ได้ โดยจะกำหนดเงื่อนไขดังนี้ คือเมื่อค่ารูปแบบเป็น 0 จะไม่
 ทำการแทรกซอลท์ลงในรหัสผ่านในตำแหน่งนั้น แต่ถ้าค่า

รูปแบบเป็น 1 จะทำการแทรกซอลท์ลงในรหัสผ่าน 1 ตัว และ
 ถ้าค่ารูปแบบเป็น 0 ติดกัน 2 ตัว จะทำการแทรกซอลท์ลงใน
 รหัสผ่าน 2 ตัว ดังตัวอย่างในตารางที่ 5 และตำแหน่งของการ
 ในแทรกซอลท์ดังแสดงในภาพที่ 2

ตารางที่ 5: ตัวอย่างการแทรกซอลท์ลงในรหัสผ่าน

รหัสผ่านตั้งต้น	password
ซอลท์	%@&03U+
ค่ารูปแบบ	01011001
ผลลัพธ์	pa%ss@w&or03U+



ภาพที่ 2: ตำแหน่งของการแทรกซอลท์

ในขั้นตอนสุดท้ายจะเป็นการแฮชค่าผลลัพธ์และจัดเก็บลงใน
 ฐานข้อมูล โดยจะทำการบันทึกค่าแฮชและค่าซอลท์ลงใน
 ฐานข้อมูลแต่จะไม่เก็บค่ารูปแบบลงไปด้วย ซึ่งทำให้สามารถ
 รักษาความปลอดภัยของรหัสผ่านและเพิ่มความซับซ้อนในการ
 โจมตีได้

4. ผลการดำเนินงาน

งานวิจัยนี้ได้ทำการพัฒนาอัลกอริทึมสำหรับการจัดเก็บ
 รหัสผ่านให้มีความปลอดภัยโดยใช้เทคนิค Dynamic salt
 selection ร่วมกับ Hash function ผู้วิจัยจึงทำการทดสอบโดย
 แบ่งออกเป็น 2 ด้านคือด้านความเร็วในการทำงานและด้านด้าน
 ความทนทานต่อการโจมตี โดยการเปรียบเทียบกับวิธีการที่มีอยู่
 ในปัจจุบัน

4.1 ผลการทดสอบด้านระยะเวลาที่ใช้ในการทำงาน

ทำการทดสอบโดยการสร้างและจัดเก็บรหัสผ่านด้วยวิธีการ
 ต่างๆที่นิยมใช้กันในปัจจุบัน [8] ลงในฐานข้อมูล ซึ่งข้อมูล
 รหัสผ่านที่นำมาใช้ในการทดลองเป็นข้อมูลรหัสผ่านจาก
 เว็บไซต์ที่ได้ทำเก็บรวบรวมว่าเป็นรหัสผ่านที่มีผู้ใช้นิยมมาก
 ที่สุด ประกอบไปด้วยรหัสผ่านทั้งหมด 10,000 รหัสผ่าน [14]
 และฐานข้อมูลที่ใช้เป็นฐานข้อมูล MySQL โดยวิธีการที่ผู้วิจัย
 ได้พัฒนาขึ้นใช้เวลา 2.48 ms ต่อรหัสผ่าน ซึ่งช้ากว่าวิธีการแบบ
 static salt 256 bit และ dynamic salt 256 bit ที่ใช้เวลาเพียง 2.04
 ms และ 2.09 ms ตามลำดับ และใช้เวลาใกล้เคียงกับวิธี

dynamic salt 80-256 bit ซึ่งมีการกำหนดให้สัมพันธ์ค่าความยาวของซอลท์เท่ากัน ซึ่งใช้เวลา 2.43 ms และทำงานรวดเร็วกว่าวิธีที่มีขนาดของซอลท์ที่มากกว่าคือ dynamic salt 512, 1024 bit และวิธีที่มีการทำแฮชฟังก์ชัน 2 รอบ dynamic salt 256 bit with 2 iteration ดังแสดงในตารางที่ 6

ตารางที่ 6: แสดงเวลาที่ใช้ในกระบวนการแฮชด้วยวิธีต่างๆ

method	hash time (ms)
static salt 256 bit	2.04
dynamic salt 64 bit	2.07
dynamic salt 256 bit	2.09
dynamic salt 512 bit	2.52
dynamic salt 1024 bit	2.56
dynamic salt 80-256 bit	2.43
dynamic salt 256 bit with 2 iteration	2.51
dynamic salt selection *	2.48

* คือวิธีที่นำเสนอ

4.2 ผลการทดสอบด้านความทนทานต่อการโจมตี

โดยใช้วิธีการจำลองสถานการณ์การโจมตี เมื่อระบบหรือแอปพลิเคชันถูกผู้ไม่หวังดีโจมตีระบบจนสามารถได้ฐานข้อมูลผู้ใช้ของระบบทั้งหมดไปได้ [1], [15] ซึ่งในฐานข้อมูลนี้จะประกอบไปด้วยข้อมูลที่เป็น ชื่อผู้ใช้ (username) ค่าซอลท์ (salt) และค่าแฮชของรหัสผ่าน (password hash) ซึ่งหลังจากผู้โจมตีได้ฐานข้อมูลรหัสผ่านมาแล้ว รหัสผ่านที่ได้มานั้นเป็นเพียงค่าแฮชซึ่งยังไม่สามารถนำไปใช้งานได้ ผู้โจมตีจึงมีความจำเป็นที่จะต้องทำการโจมตีค่าแฮช (hash password) เหล่านี้ให้กลายเป็นคำรหัสผ่าน (plaintext password) เสียก่อน ซึ่งในงานวิจัยนี้ได้ใช้โปรแกรม hashcat ซึ่งเป็น โปรแกรมที่มีชื่อเสียงและได้รับความนิยมทางด้านการกู้คืนรหัสผ่าน [16] ในการโจมตีค่าแฮช

รหัสผ่านที่นำมาทดสอบในงานวิจัยนี้จะใช้รายการรหัสผ่านที่รวบรวมโดย SplashData [17] ซึ่งเป็นรหัสผ่านที่ถูกใช้งานมากที่สุดและถือเป็นรหัสผ่านที่แย่ที่สุดของปี 2013 มาทำการผ่านแฮชฟังก์ชันด้วยวิธีการต่าง ๆ รวมถึงเทคนิคที่ผู้วิจัยได้พัฒนาขึ้น แล้วทำการจัดเก็บลงในฐานข้อมูล MySQL เพื่อใช้ในการทดสอบการโจมตี ซึ่งรายการของรหัสผ่าน 25 รหัสที่ใช้ในการทดลองนี้แสดงในภาพที่ 3

Rank	Password	Rank	Password
1	123456	14	letmein
2	password	15	photoshop
3	12345678	16	1234
4	qwerty	17	monkey
5	abc123	18	shadow
6	123456789	19	sunshine
7	111111	20	12345
8	1234567	21	password1
9	iloveyou	22	princess
10	adobe123	23	azerty
11	123123	24	trustno1
12	admin	25	000000
13	1234567890		

ภาพที่ 3: แสดงรายการรหัสผ่านที่มีผู้ใช้งานมากที่สุดในปี 2013

ทำการโจมตีรหัสผ่าน ด้วยโปรแกรม hashcat ผลการทดสอบพบว่าการจัดเก็บด้วยวิธีทั่วไปสามารถโจมตีรหัสผ่านได้ 23 รหัสจากทั้งหมด 25 รหัส หรือ 92% จากรหัสผ่านทั้งหมด ส่วนวิธีการที่ผู้วิจัยนำเสนอ นั้นไม่สามารถโจมตีได้เลย ดังแสดงในตารางที่ 7

ตารางที่ 7: แสดงผลการโจมตีค่าแฮชด้วยโปรแกรม hashcat

วิธีการจัดเก็บรหัสผ่าน	รูปแบบ	success
no salt	md5(\$password)	92%
multiple iteration	md5(md5(md5(\$password)))	92%
fix salt	md5(\$password.fixsalt)	92%
dynamic salt	md5(\$password.dynamicsalt)	92%
dynamic salt selection*	md5(\$password.selectionsalt)	0%

* คือวิธีที่นำเสนอ

5. สรุป

งานวิจัยนี้ได้เสนออัลกอริทึมสำหรับการจัดเก็บรหัสผ่านอย่างปลอดภัยโดยใช้เทคนิคการปรับค่าซอลท์ที่เหมาะสมร่วมกับแฮชฟังก์ชัน ผลจากการทดลองพบว่าในสถานการณ์จำลองการโจมตีที่ระบบถูกโจรกรรมฐานข้อมูลรหัสผ่าน (username, salt, password hash) ออกไป ยังคงสามารถรักษาความลับของรหัสผ่านเหล่านั้นไว้ได้ และระยะเวลาที่ใช้ในการทำงานใกล้เคียงกับแบบใช้วิธีการสุ่ม โดยทั่วไปในกรณีที่ซอลท์ที่มีขนาดความยาวเท่ากันและยังพบว่าการนำรหัสผ่าน ไปผ่าน

กระบวนการแฮชฟังก์ชันหลายๆรอบไม่สามารถทำให้รักษาความลับของรหัสผ่านให้ดีขึ้น และอัลกอริทึมนี้ยังมีการสร้างค่ารูปแบบในการแทรกซอลท์ลงในรหัสผ่าน ซึ่งจะเพิ่มความซับซ้อนในการโจมตีอีกด้วย อย่างไรก็ตามในอนาคตสามารถพัฒนาวิธีการอื่นที่ใช้ในการสร้างค่ารูปแบบหรือข้อกำหนดที่ใช้ในการแทรกซอลท์ลงในรหัสผ่านให้มีความซับซ้อนยิ่งขึ้น ทำให้สามารถลดความเสี่ยงเมื่อถูกโจรกรรมข้อมูลได้

เอกสารอ้างอิง

- [1] H. Kumar, S. Kumar, R. Joseph, D. Kumar, S. K. Shrinarayan Singh, P. Kumar, and H. Kumar, "Rainbow table to crack password using MD5 hashing algorithm," *IEEE Conference on Information Communication Technologies (ICT)*, pp. 433–439, 2013.
- [2] A. A. Putri Ratna, P. Dewi Purnamasari, A. Shaugi, and M. Salman, "Analysis and comparison of MD5 and SHA-1 algorithm implementation in Simple-O authentication based security system," *International Conference on QiR (Quality in Research)*, pp. 99–104, 2013.
- [3] C. S. Kumari and M. D. Rani, "Hacking resistance protocol for securing passwords using personal device," *7th International Conference on Intelligent Systems and Control (ISCO)*, pp. 458–463, 2013.
- [4] Q. Wang and Z. Qin, "Stronger user authentication for web browser," *3rd International Conference 2010 on Advanced Computer Theory and Engineering (ICACTE)*, vol. 5, pp. 539–543, 2010.
- [5] P. Li, Y. Sui, and H. Yang, "The parallel computation in one-way hash function designing," *International Conference on Computer, Mechatronics, Control and Electronic Engineering (CMCE)*, pp. 189–192, 2010.
- [6] X. Zheng and J. Jin, "Research for the application and safety of MD5 algorithm in password authentication," *9th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, pp. 2216–2219, 2012.
- [7] S. K. Sood, A. K. Sarje, and K. Singh, "Cryptanalysis of password authentication schemes: Current status and key issues," *Proceeding of International Conference on Methods and Models in Computer Science, ICM2CS*, pp. 1–7, 2009.
- [8] P. Gauravaram, "Security Analysis of salt|password Hashes," *International Conference on Advanced Computer Science Applications and Technologies (ACSAT)*, pp. 25–30, 2012.
- [9] S. Boonkrong, "Security of Passwords" *Journal of Information Technology* Vol. 8, No. 2, July – December 2012 pp. 112-117.
- [10] K. Theocharoulis, I. Papaefstathiou, and C. Manifavas, "Implementing Rainbow Tables in High-End FPGAs for Super-Fast Password Cracking," *International Conference on Field Programmable Logic and Applications (FPL)*, pp. 145–150, 2010.
- [11] W. Ma, J. Campbell, D. Tran, and D. Kleeman, "Password Entropy and Password Quality," *4th International Conference on Network and System Security (NSS)*, pp. 583–587, 2010.
- [12] X. Zheng and J. Jin, "Research for the application and safety of MD5 algorithm in password authentication," *9th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, pp. 2216–2219, 2012.
- [13] Jeff Atwood. "Speed Hashing" [Online] 2012 Availablefrom: <http://www.codinghorror.com/blog/2012/04/speed-hashing.html> [2013, July 1].
- [14] Mark, B. "Artificial intelligence" [Online] Availablefrom: <https://xato.net/passwords/more-top-worst-passwords/#.UwXIVYV1BRI> [2013, July 10]
- [15] Whitney, L. "Millions of LinkedIn passwords reportedly leaked online" [Online] 2012 Availablefrom: http://news.cnet.com/8301-1009_3-57448079-83/millions-of-linkedin-passwords-reportedly-leaked-online/ [2013, July 1]
- [16] Kali "How to crack passwords using Hashcat – The Visual Guide" [Online] 2010 Availablefrom: <http://uwntesis.wordpress.com/2013/08/07/kali-how-to-crack-passwords-using-hashcat/> [2013, July 1]
- [17] Kevin, D. "The 2013 list of worst passwords" [Online] 2013 Availablefrom: <http://splashdata.com/press/worstpasswords2013.htm> [2013, October 25]

การป้องกันการปลอมแปลงสิทธิ์ใช้งานระบบเครือข่าย ด้วยค่า UUID และ MAC Address Prevention of Fake Authentication on Network System by UUID and MAC Address Audit

สง่าวุฒิ โปธา (Sangavood Pota)¹ และ ชัชวิน นามมัน (Chatchawin Namman)²
ภาควิชาคณิตศาสตร์สถิติ และคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยอุบลราชธานี
¹tengtong_tongteng@hotmail.com, ²scchatna@ubu.ac.th

บทคัดย่อ

งานวิจัยนี้ เป็นการพัฒนากระบวนการเพิ่มความปลอดภัยให้กับระบบเครือข่ายที่ใช้การแจกจ่ายชุดหมายเลขไอพีแอดเดรสแบบไดนามิก โดยผู้วิจัยได้ใช้วิธีการกรองค่าแมคแอดเดรสร่วมกับการตรวจสอบค่าหมายเลขยูยูไอดีของเครื่องลูกข่าย เพื่อพิสูจน์ทราบตัวตนก่อนใช้งานระบบเครือข่าย ระบบถูกพัฒนาขึ้นด้วยภาษา C# ผลจากการทดสอบและประเมินประสิทธิภาพของระบบในด้านความถูกต้องของระบบ ด้านการใช้งานระบบ และด้านความปลอดภัยของระบบจากผู้เชี่ยวชาญ จำนวน 5 คน พบว่าประสิทธิภาพโดยรวมของระบบมีค่าเฉลี่ยเท่ากับ 4.48 และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.49 จึงสรุปได้ว่าระบบที่พัฒนาขึ้นมีประสิทธิภาพอยู่ในระดับดี และสามารถนำไปประยุกต์ใช้งานต่อไปได้

คำสำคัญ: ไดนามิก, การกรองค่าแมคแอดเดรส, ยูยูไอดี, การพิสูจน์ทราบตัวตน

Abstract

This research was developed an alternative approach to enhancing a network security distributing an IP address by DHCP. We used MAC address filtering with client's UUID as a way to an authentication before accessing on the networks. The system was implemented by using C#. It was evaluated in term of accurate and security from five experts. The result showed that the average efficiency was 4.48 with standard deviation at 0.49. It can be concluded that the proposed system performance is at a good level and can be beneficial for further applications.

Keywords: DHCP, MAC Address Filtering, UUID, Authentication

1. บทนำ

รูปแบบในการเชื่อมต่อระบบเครือข่ายที่หลากหลายวิธี และวิธีการในการแจกจ่ายหมายเลขไอพีแอดเดรสที่นิยมใช้ในปัจจุบันเป็นในลักษณะของไดนามิก [1] ซึ่งจะทำให้การแจกจ่ายชุดไอพีแอดเดรสแบบอัตโนมัติให้แก่เครื่องลูกข่ายที่ร้องขอเข้ามาใช้บริการ ปัญหาในเรื่องการแจกจ่ายชุดหมายเลขไอพีแอดเดรสในลักษณะดังกล่าวที่ขาดการกำหนดนโยบายด้านความปลอดภัยระบบเครือข่าย ปัจจุบันมีการนำเทคนิคการกรองผู้ใช้งานระบบโดยการตรวจสอบค่าแมคแอดเดรสของเครื่องลูกข่ายก่อนใช้งานระบบเครือข่าย ปัญหาที่พบจากวิธีการดังกล่าวที่อาจเป็นช่องโหว่ให้ผู้ประสงค์ร้ายทำการดักจับข้อมูลและแอบอ้างค่าแมคแอดเดรสเพื่อทำการปลอมแปลงสิทธิ์ใช้งานระบบเครือข่ายได้ ดังนั้นจึงได้นำเทคนิคในเรื่องของการกรองผู้ใช้งานโดยการตรวจสอบค่าชุดหมายเลขยูยูไอดีและแมคแอดเดรสเพื่อใช้พิสูจน์ทราบตัวตนของผู้ใช้ก่อนใช้งานระบบเครือข่ายเพื่อนำไปใช้ในการเพิ่มความปลอดภัยระบบเครือข่ายภายในองค์กรให้ดียิ่งขึ้น โดยในบทความนี้จะกล่าวถึงงานวิจัยที่เกี่ยวข้องวิธีการดำเนินการวิจัย ผลการดำเนินการวิจัย และสรุปตามลำดับ

2. งานวิจัยที่เกี่ยวข้อง

จากการศึกษางานวิจัยที่เกี่ยวข้องกับการแจกจ่ายไอพีแอดเดรสแบบไดนามิก และการพิสูจน์ทราบตัวตนผู้ใช้งาน โดยสามารถสรุปได้ดังต่อไปนี้

ระบบตรวจสอบและป้องกันการเข้าถึงเครือข่ายของคอมพิวเตอร์แปลกปลอมภายในเครือข่ายเดียว [2] โดยนำ

หลักการการทำงานของ ARP Protocol และใช้วิธีการตรวจสอบค่าหมายเลขไอพีแอดเดรสและแมคแอดเดรส แล้วทำการส่งค่าไปตรวจสอบค่าในเชิงเปรียบเทียบว่าถูกต้องและตรงกับค่าที่มีอยู่ในฐานข้อมูลจึงอนุญาตให้เข้าใช้งานเครือข่ายได้ จากการวิเคราะห์สรุปได้ว่าเทคนิคดังกล่าวยังมีความเสี่ยงในเรื่องของการถูกดักจับข้อมูลภายในเครือข่ายโดยการปลอมแปลงค่าแมคแอดเดรสและปลอมแปลงสิทธิ์เข้าใช้งานเครือข่ายได้

การวิเคราะห์ข้อมูลของจุดให้บริการเครือข่ายแบบไร้สายภายใต้การทำงานของดีเอชซีพีโปรโตคอล [3] โดยพัฒนารูปแบบเพื่อใช้ในการวิเคราะห์ข้อมูลของจุดให้บริการเครือข่ายแบบไร้สายภายใต้การทำงานของดีเอชซีพีโปรโตคอล ร่วมกับการระบุตัวตนในเรื่องของใบรับรองดิจิทัลโดยใช้กุญแจสาธารณะ จากการวิเคราะห์พบว่าการทำงานภายใต้ของดีเอชซีพีโปรโตคอลที่ขาดต่อการกำหนดนโยบายด้านความปลอดภัยในเรื่องของการระบุตัวตนเพื่อพิสูจน์ทราบตัวตนผู้ใช้งานเครือข่าย

กระบวนการเพิ่มความปลอดภัยของดีเอชซีพีโปรโตคอลร่วมกับใบรับรองดิจิทัล [4] โดยพัฒนารูปแบบการติดต่อสื่อสารภายใต้ดีเอชซีพีโปรโตคอล โดยศึกษากระบวนการทำงานของดีเอชซีพีเซิร์ฟเวอร์และดีเอชซีพีไคลเอนต์ในการส่งข้อความโต้ตอบระหว่างเครื่องลูกข่ายกับเครื่องแม่ข่าย โดยการเปรียบเทียบการโต้ตอบภายใต้แนวคิดในเรื่องความปลอดภัยในชุดดีซีพีโปรโตคอล จากการวิเคราะห์สรุปได้ว่าเทคนิคดังกล่าวอาจถูกดักจับข้อมูลการเชื่อมต่อภายในเครือข่ายเนื่องจากการทำงานภายใต้ดีเอชซีพีโปรโตคอลที่มีการส่งข้อมูลแบบกระจายและไม่รับประกันความถูกต้องในการส่งข้อมูลในแต่ละครั้ง

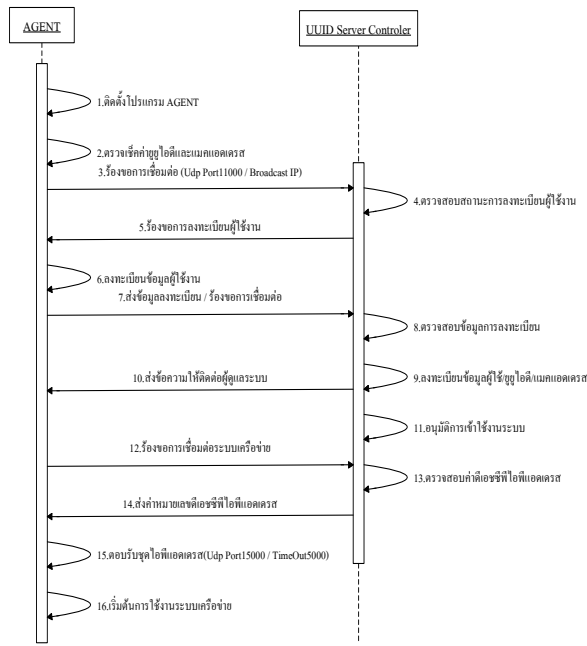
รูปแบบของการระบุตัวตนของออบเจกต์หรือเอนทิตีในรูปแบบของ GUID [5] โดยทำการศึกษารูปแบบของการระบุตัวตนของออบเจกต์หรือเอนทิตีในระบบเครือข่ายที่มีความซับซ้อนและขาดต่อการเข้าถึงข้อมูล จากการวิเคราะห์สามารถสรุปได้ว่างานวิจัยนี้มีประโยชน์ในเรื่องของการระบุตัวตนของออบเจกต์ที่ให้ค่าที่ไม่ซ้ำกัน

จากการศึกษางานวิจัยที่เกี่ยวข้อง ที่ได้ทำการศึกษาพบว่า ยังไม่มีการพัฒนารูปแบบการระบุตัวตน โดยใช้ชุดหมายเลขแมคแอดเดรสร่วมกับค่ายูยูไอดีเพื่อใช้ในเรื่องของการพิสูจน์

ทราบตัวตนของเครื่องลูกข่าย โดยค่าชุดตัวเลขยูยูไอดี [6] ซึ่งเป็นค่าที่ถูกกำหนดภายใต้มาตรฐานของ IETF และการกำหนดรูปแบบโดยใช้ RFC 4122 UUID Namespace ซึ่งใช้ระบุนค่าที่ไม่ซ้ำกันของออบเจกต์โดยค่าที่ได้จะเป็นชุดตัวเลขฐาน 16 ที่มีขนาด 128 บิต โดยการใช้เทคนิควิธีการแบบเวลาประทับ ซึ่งให้ค่าชุดตัวเลขที่มีความซับซ้อน ยกต่อการเข้าถึง และการให้ค่าที่ไม่ซ้ำกัน ซึ่งมีความน่าเชื่อถือในเรื่องของการระบุความเป็นตัวตนของออบเจกต์ภายในเครือข่าย จึงทำให้ผู้วิจัยมีความสนใจเป็นอย่างมาก ในการพัฒนาเทคนิควิธีการและรูปแบบการระบุตัวตนโดยการตรวจสอบค่าแมคแอดเดรสร่วมกับค่ายูยูไอดีของเครื่องลูกข่าย เพื่อใช้ในการพิสูจน์ทราบตัวตนผู้ใช้งานก่อนเข้าใช้งานระบบเครือข่าย

3. วิธีการดำเนินการวิจัย

การพัฒนากระบวนการป้องกันการปลอมแปลงสิทธิ์เข้าใช้งานระบบเครือข่าย ด้วยค่ายูยูไอดีและแมคแอดเดรส ผู้วิจัยได้พัฒนาระบบด้วยชุดโปรแกรมประยุกต์ Visual C#.NET และใช้ชุดโปรแกรมการจำลองระบบปฏิบัติการเสมือน VMWare Workstation ซึ่งประกอบไปด้วยระบบปฏิบัติการวินโดวส์เซิร์ฟเวอร์ 2003 จำนวน 1 เครื่อง เพื่อทำหน้าที่เป็นเครื่องดีเอชซีพีเซิร์ฟเวอร์ ระบบปฏิบัติการวินโดวส์เอ็กพี จำนวน 3 เครื่อง เพื่อทำหน้าที่เป็นดีเอชซีพีไคลเอนต์ และระบบปฏิบัติการวินโดวส์เอ็กพี จำนวน 1 เครื่อง เพื่อใช้ทดสอบการปลอมแปลงสิทธิ์เข้าใช้งานเครือข่าย ซึ่งการวิเคราะห์ระบบในการส่งข้อความโต้ตอบระหว่างเครื่องลูกข่ายและเครื่องแม่ข่าย ภายใต้การแจกจ่ายชุดหมายเลขไอพีแอดเดรสแบบดีเอชซีพีผ่านทางยูดีพีโปรโตคอล แสดงดังภาพที่ 1



ภาพที่ 1: กระบวนการทำงานของระบบ

จากภาพที่ 1 แสดงกระบวนการทำงานทั้งหมดของระบบ ซึ่งมีดังต่อไปนี้ 1) เริ่มต้นโดยผู้ใช้งานทำการติดตั้งโปรแกรม AGENT ที่เครื่องลูกข่าย 2) โปรแกรมทำการค้นหาค่ายูไอดีและแมคแอดเดรสของเครื่องลูกข่าย 3) เครื่องลูกข่ายร้องขอการเชื่อมต่อออกไปยังเครื่องข่ายผ่านยูดีพีพอร์ตหมายเลข 11000 ในลักษณะ broadcast ไอพีแอดเดรส 4) โปรแกรม UUID Server Controller ที่เครื่องแม่ข่ายตอบรับการเชื่อมต่อและตรวจสอบสถานะการลงทะเบียนเข้าใช้งานระบบครั้งแรกของผู้ใช้ 5) เครื่องแม่ข่ายร้องขอข้อมูลการลงทะเบียนผู้ใช้งานกลับไปยังเครื่องลูกข่าย 6) ผู้ใช้งานกรอกข้อมูลระบบซึ่งได้แก่หมายเลขไอดีการ์ด 13 หลัก และพาสเวิร์ดที่ได้รับจากผู้ดูแลระบบเพื่อใช้ลงทะเบียน 7) โปรแกรมส่งค่าข้อมูลกลับไปยังเครื่องแม่ข่ายเพื่อร้องขอการลงทะเบียนผู้ใช้งาน 8) เครื่องแม่ข่ายตรวจสอบความถูกต้องของข้อมูลการลงทะเบียน 9) ระบบลงทะเบียนข้อมูลผู้ใช้งานพร้อมทั้งบันทึกค่าชุดหมายเลขยูไอดีและแมคแอดเดรส 10) ในครั้งแรกของการร้องขอการเชื่อมต่อ ระบบจะส่งข้อความแจ้งเตือนกลับไปให้เครื่องลูกข่ายเพื่อให้ผู้ใช้ติดต่อผู้ดูแลระบบเพื่อให้ทำการอนุมัติเข้าใช้งานเครื่องข่าย ซึ่งเป็นการยืนยันตัวตนผู้ใช้งาน 11) ผู้ดูแลระบบทำการอนุมัติให้เครื่องลูกข่ายสามารถเข้าใช้งานระบบได้ 12) ผู้ใช้งานทำการ ร้องขอการเชื่อมต่อเครื่องข่ายไปยังเครื่องแม่ข่ายอีกครั้ง

13) เครื่องแม่ข่ายทำการกำหนดค่าชุดไอพีแอดเดรสภายในระบบที่ว่างและไม่มีการใช้งาน 14) ระบบส่งค่าชุดไอพีแอดเดรสไปให้เครื่องลูกข่าย 15) เครื่องลูกข่ายตอบรับชุดไอพีแอดเดรสผ่านทางยูดีพีพอร์ตหมายเลข 15000 และกำหนดเวลาในการตอบรับชุดข้อมูลเท่ากับ 5000 16) เครื่องลูกข่ายเริ่มต้นใช้งานชุดหมายเลขไอพีแอดเดรส และสามารถเข้าใช้งานระบบเครื่องข่ายได้

4. ผลการดำเนินการวิจัย

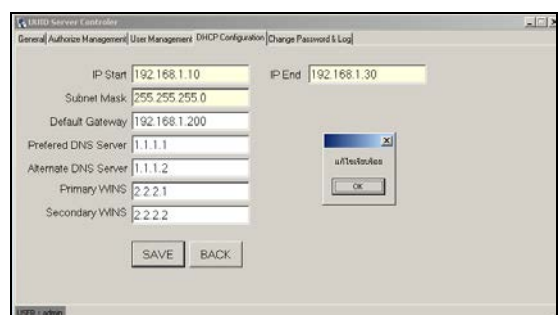
ผลการพัฒนาระบบการป้องกันการปลอมแปลงสิทธิ์เข้าใช้งานระบบเครื่องข่าย ด้วยค่ายูไอดีและแมคแอดเดรส ประกอบด้วยการทำงานในส่วนต่าง ๆ ดังต่อไปนี้

4.1 ผลการพัฒนาระบบในส่วนของเครื่องแม่ข่าย

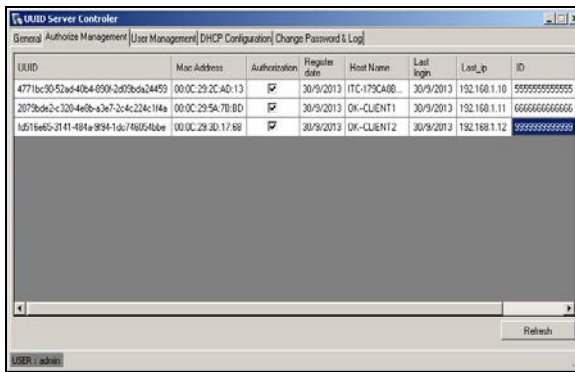
โดยการพัฒนาโปรแกรม UUID Server Controller ประกอบด้วยเมนูหลักของการทำงาน ได้แก่ การเพิ่มข้อมูลผู้ใช้งานระบบ (ภาพที่ 2) การกำหนดขอบเขตของการแจกจ่ายไอพีแอดเดรส (ภาพที่ 3) การอนุมัติเข้าใช้งานระบบ (ภาพที่ 4) และการบันทึกล็อกไฟล์ (ภาพที่ 5) ตามลำดับ



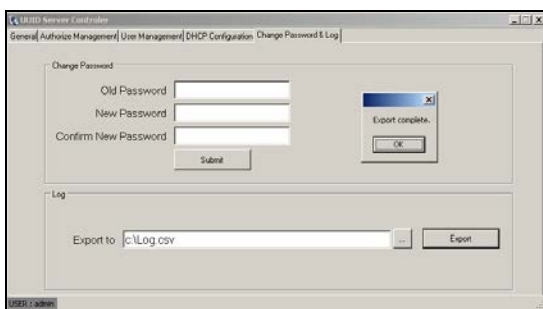
ภาพที่ 2: เพิ่มข้อมูลผู้ใช้งานระบบเครื่องข่าย



ภาพที่ 2: การกำหนดขอบเขตการแจกจ่ายชุดไอพีแอดเดรส



ภาพที่ 3: การอนุมัติให้เครื่องลูกข่ายเข้าใช้งานเครือข่าย



ภาพที่ 4: การทึกรหัสผ่านผู้ใช้การเข้าใช้งานเครือข่าย



ภาพที่ 6: การเปลี่ยนรหัสผ่านของผู้ใช้งานระบบ



ภาพที่ 7: แสดงสถานะพร้อมใช้งานระบบเครือข่าย

4.2 ผลการพัฒนาในระบบในส่วนเครื่องลูกข่าย

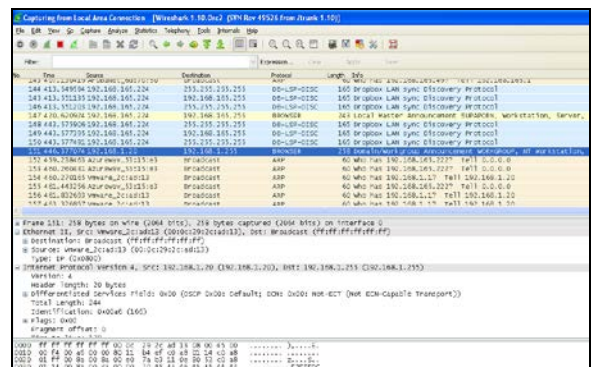
ในส่วนเครื่องลูกข่ายพัฒนาโปรแกรม AGENT ซึ่งสามารถส่งค่าการลงทะเบียนผู้ใช้ (ภาพที่ 5) การเปลี่ยนรหัสผ่านผู้ใช้งานระบบ (ภาพที่ 6) การร้องขอและการยกเลิกการเชื่อมต่อระบบเครือข่าย (ภาพที่ 7)



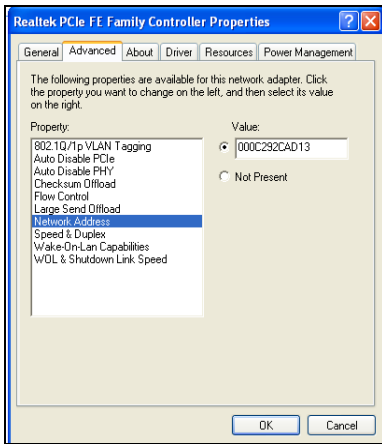
ภาพที่ 5: การลงทะเบียนผู้ใช้งานเครือข่าย

4.3 ผลการทดสอบการปลอมแปลงสิทธิ์เข้าใช้งานเครือข่าย

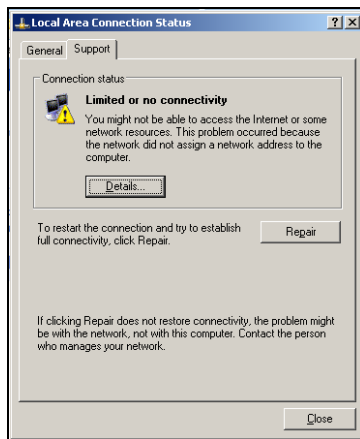
กระบวนการทดสอบโดยใช้วิธีการดักจับข้อมูลภายในระบบเครือข่าย (ภาพที่ 8) และทำการแอบอ้างค่าหมายเลขแมคแอดเดรสของเครื่องลูกข่ายเพื่อใช้ปลอมแปลงสิทธิ์เข้าใช้งานระบบเครือข่าย (ภาพที่ 9) ซึ่งผลการทดสอบพบว่าไม่สามารถแอบอ้างค่าหมายเลขแมคแอดเดรสเพื่อใช้ปลอมแปลงสิทธิ์เข้าใช้งานระบบเครือข่ายได้ (ภาพที่ 10)



ภาพที่ 8: แสดงการดักจับข้อมูลโดยใช้โปรแกรม Wireshark



ภาพที่ 9: การแอบอ้างค่าแมคแอดเดรสเพื่อใช้ปลอมแปลงสิทธิ์



ภาพที่ 10: สถานะการเชื่อมต่อที่ไม่สามารถเข้าใช้งานเครือข่ายได้

4.4 ผลการประเมินระบบ

หลังจากที่ได้พัฒนาระบบและการทดสอบระบบแล้ว ขั้นตอนการประเมินผล โดยทำการทดสอบใช้งานระบบ ทั้ง 3 ด้าน ซึ่งประกอบด้วย การประเมินด้านความถูกต้องของระบบ การประเมินด้านการใช้งานระบบ และการประเมินด้านความปลอดภัยของระบบ โดยใช้แบบสอบถามจากผู้เชี่ยวชาญ จำนวน 5 ท่าน ผลการประเมินผลด้านต่าง ๆ รวมทั้งการประเมินผลโดยรวมของระบบ แสดงดังตารางที่ 1

ตารางที่ 1: สรุปผลการประเมินระบบด้านต่าง ๆ

รายการประเมิน	ผลความคิดเห็นของผู้ประเมิน		
	\bar{x}	SD	ความหมาย
1.การประเมินด้านความถูกต้องของระบบ	4.57	0.44	ดีมาก
2.การประเมินด้านการใช้งานระบบ	4.40	0.52	ดี
3.การประเมินด้านความปลอดภัยของระบบ	4.47	0.51	ดีมาก
เฉลี่ย	4.48	0.49	ดี

จากตารางที่ 1 แสดงผลการประเมินระบบในด้านต่าง ๆ ซึ่งประกอบด้วย ผลการประเมินด้านความถูกต้องของระบบ มีค่าเฉลี่ยเท่ากับ 4.48 และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.44 ผลการประเมินด้านการใช้งานระบบมีค่าเฉลี่ยเท่ากับ 4.40 และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.52 ผลการประเมินด้านความปลอดภัยของระบบมีค่าเฉลี่ยเท่ากับ 4.47 และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.51 ซึ่งเมื่อนำคะแนนเฉลี่ยของแต่ละหัวข้อมาผ่านระเบียบวิธีการทางสถิติเพื่อหาค่าเฉลี่ย พบว่าค่าเฉลี่ยโดยรวมเท่ากับ 4.48 และส่วนเบี่ยงเบนมาตรฐานโดยรวมเท่ากับ 0.49 ซึ่งแสดงว่าการกระจายของการให้คะแนนที่ไม่แตกต่างกัน จึงสรุปได้ว่าระบบที่พัฒนาขึ้นมีประสิทธิภาพอยู่ในระดับดี

5. สรุป

ในงานวิจัยนี้ผู้วิจัยได้ออกแบบและพัฒนากระบวนการเพื่อใช้ในการติดต่อสื่อสารระหว่างเครื่องลูกข่ายและเครื่องแม่ข่าย โดยพัฒนาระบบการป้องกันการปลอมแปลงสิทธิ์เข้าใช้งานระบบเครือข่ายด้วยค่ายูยูไอดีและแมคแอดเดรส ผู้วิจัยได้พัฒนาระบบด้วยชุดโปรแกรมประยุกต์ Visual C#.NET และใช้โปรแกรมในการจำลองระบบปฏิบัติการเสมือน VMWare Workstation ในการสร้างแบบจำลองระบบปฏิบัติการเสมือนจริงเพื่อใช้ในการทดสอบ

ในสภาพแวดล้อมระบบเครือข่ายในลักษณะการเชื่อมต่อแบบใช้สายแลนและการแจกจ่ายชุดไอพีแอดเดรสแบบดีเอชซีพี

ซึ่งผลการพัฒนาระบบการป้องกันการปลอมแปลงสิทธิ์ใช้งานระบบเครือข่าย ด้วยคำยูยูไอดีและแมคแอดเดรสเป็นการพัฒนากระบวนการในเรื่องของการพิสูจน์ทราบตัวตนของผู้ใช้งานเครือข่ายที่มีความซับซ้อนและยากต่อการเข้าถึงชุดข้อมูลภายใต้ระบบเครือข่ายที่ใช้การแจกจ่ายชุดหมายเลขไอพีแอดเดรสแบบดีเอชซีพีที่ทำงานในลักษณะของการติดต่อสื่อสารแบบคอนเน็คชันเลส ซึ่งไม่มีการรับประกันความถูกต้องของการรับส่งข้อมูล ซึ่งผลการพัฒนาระบบประกอบด้วยส่วนของเครื่องแม่ข่ายซึ่งสามารถกำหนดขอบเขตการแจกจ่ายชุดหมายเลขไอพีแอดเดรส การเพิ่มข้อมูลผู้ใช้งาน การอนุมัติการเข้าใช้งานเครือข่าย และการเก็บบันทึกข้อมูลล็อกไฟล์ ในส่วนของเครื่องลูกข่าย สามารถทำการลงทะเบียนเข้าใช้งานระบบ การเปลี่ยนรหัสผ่านของผู้ใช้งาน การร้องขอการเชื่อมต่อและการยกเลิกการใช้งานระบบเครือข่ายได้ รวมทั้งผลการประเมินประสิทธิภาพของระบบ ซึ่งประกอบด้วย การทดสอบระบบด้านความถูกต้องของระบบ ด้านการใช้งานระบบ และด้านความปลอดภัยของระบบ โดยผู้เชี่ยวชาญจำนวน 5 ท่าน ซึ่งผลการประเมินผลโดยรวมทั้ง 3 ด้าน พบว่าค่าเฉลี่ยโดยรวมเท่ากับ 4.48 และส่วนเบี่ยงเบนมาตรฐานโดยรวมเท่ากับ 0.49 แสดงว่าการกระจายของการให้คะแนนที่ไม่แตกต่างกัน จึงสรุปได้ว่าระบบที่พัฒนาขึ้นมีประสิทธิภาพอยู่ในระดับดี และนำไปประยุกต์ใช้งานด้านความปลอดภัยระบบเครือข่ายภายในองค์กรต่อไปได้

เอกสารอ้างอิง

- [1] Droms R. Dynamic Host Configuration Protocol. [online]. [สืบค้นวันที่ 18 เมษายน 2556]. <http://www.ietf.org/rfc/rfc2131.txt>
- [2] สุริยะ พินิจการ, “ระบบตรวจสอบและป้องกันการเข้าถึงเครือข่ายของคอมพิวเตอร์แปลกปลอมภายในเครือข่ายเดียว” สารนิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2553.
- [3] Tuomas Aura, Michael Roe and Steve J. Murdoch, “Security Network Location Awareness with Authenticated DHCP”, *IEEE Communications Networks and the Workshop Third International Conference SecureComm*, vol. 10, no. 1 pp. 391-402, 2007.
- [4] สุรการ ดวงผาสุข และคณะ, “Design and Implementation of Improved Security Protocols for DHCP Using Digital Certificates,” *Networks (ICON) 2011 17th IEEE International Conference*, vol. 10, no. 1, pp. 287-292, 2011.
- [5] Christof Lutteroth and Gerald Weber, “Efficient Use of GUIDs”, *2008 Ninth International Conference on Parallel and Distributed Computing, Applications and Technologies*, vol. 10, no. 1 pp. 115-120, 2008.
- [6] IETF. RFC4122 A Universally Unique IDentifier (UUID) URN Namespace, [สืบค้นวันที่ 5 พฤษภาคม 2556]. <http://www.ietf.org/rfc/rfc4122.txt>

การวิเคราะห์โครงข่ายถนนโดยใช้ทฤษฎีกราฟ

กรณีศึกษา โครงข่ายถนนในจังหวัดนนทบุรี

The Analysis of Street Networks by Graph Theory: The Case Study of Street in The Nonthaburi Province

องอาจ อุ่นอนันต์ (Aongart Aun-a-nan)¹ วรวิทย์ ยิ้มแย้ม (Worawut Yimyam)² และสุนันทา สดสี (Sunantha Sodsee)³

^{1,3}ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

²สาขาวิชาคอมพิวเตอร์ธุรกิจ คณะวิทยาการจัดการ มหาวิทยาลัยราชภัฏเพชรบุรี

¹aidedecampc31020@hotmail.co.uk, ²worawut_yimyam@hotmail.com, ³sunanthas@kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์ในการวิเคราะห์การเชื่อมต่อของโครงข่ายถนนภายในจังหวัดนนทบุรีเพื่อค้นหาทางแยกที่มีการจราจรคับคั่งและคำนวณหาอัตราการรองรับปริมาณรถของแต่ละทางแยกโดยใช้ทฤษฎีกราฟ (Graph Theory) ในการสร้างแบบจำลองทางโครงข่ายและวิเคราะห์ค่า Betweenness Centrality (BC) ในการหาทางแยกที่มีการจราจรคับคั่ง และค่าประสิทธิภาพ (Capability: C) ในการหาอัตราการรองรับปริมาณรถ นำค่าทั้งสองมาสร้างแบบจำลองทางโครงข่ายแล้วเปรียบเทียบกัน ซึ่งผลการวิเคราะห์พบว่าทางแยกที่มีค่า BC มากที่สุดคือ จุด NB28 ทางแยกถนนราชพฤกษ์ตัดถนนรัตนาธิเบศร์ (BC = 0.318) ส่งผลให้มีความหนาแน่นของการจราจรสูงและทางแยกที่มีค่าประสิทธิภาพมากที่สุดคือ จุด NB13 ทางแยกถนนกาญจนาตัดถนนนครอินทร์ (C = 1.000) จากการศึกษาพบว่าทางแยกที่มีความหนาแน่นของการจราจรสูงแต่มีอัตราการรองรับปริมาณรถได้น้อยเป็นสาเหตุทำให้เกิดปัญหาการจราจรติดขัด

คำสำคัญ: ทฤษฎีกราฟ โครงข่ายถนน ศูนย์กลาง ประสิทธิภาพ

Abstract

This research aims to analyze street networks in Nonthaburi province for searching the high traffic of crossroads and calculating flow rate of traffic. Graph Theory and Complex Network were applied to create the network model. Betweenness Centrality (BC) and Capability (C) were used to find heavy crossroad traffics and flow rate of traffic. The results showed

Ratchaphruek-Rattanathibet crossroad (NB28) got the highest BC value, which was $BC = 0.318$ presenting highest traffic. On the other hand, the highest Capability value (C), Kanchanaphisek-Nakhonin crossroad (NB13) got $C = 1.00$, which is indicating the highest flow rate of traffic. It concluded that the crossroads with high BC and low C are causing to the jammed traffic.

Keyword: Graph Theory, Street Network, Betweenness Centrality, Capability.

1. บทนำ

ในปัจจุบันการจราจรในเมืองขนาดใหญ่ มีโครงข่ายถนนที่ซับซ้อน โดยถนนเหล่านี้ได้สร้างการเชื่อมโยงกันของประชากร สินค้า บริการ และเงินทุนให้สามารถเคลื่อนย้ายภายในเมืองและระหว่างเมืองได้อย่างสะดวก [1] แต่ปริมาณของรถยนต์เพิ่มมากขึ้น ส่งผลให้เกิดปัญหาการจราจรคับคั่งตามทางแยก แม้มีการบริหารจัดการสัญญาณไฟจราจร [2] ตรวจสอบจราจร และป้ายจราจรอัจฉริยะ ช่วยเหลือในทางแยกต่างๆ แต่ก็ยังเกิดปัญหาการจราจรคับคั่งอยู่ ถ้าทำการวิเคราะห์ทางแยกที่มีการจราจรคับคั่งได้ จะช่วยจัดการวางแผนในด้านการคลี่คลายอุบัติเหตุ การปิดถนน และช่วยในการคาดคะเนความหนาแน่นของการจราจรในอนาคตที่เกิดขึ้นอย่างถูกต้องมากยิ่งขึ้น

จากเหตุผลดังกล่าวงานวิจัยนี้จึงมีวัตถุประสงค์เพื่อศึกษาพฤติกรรมของการเชื่อมต่อของโครงข่ายถนน เพื่อวิเคราะห์หาทางแยกที่มีการจราจรคับคั่ง และหาความสามารถในการรองรับปริมาณรถ โดยมีการประยุกต์ใช้ทฤษฎีกราฟ [3] และการ

วิเคราะห์ Complex Networks [4] มาช่วยในการสร้างแบบจำลองโครงข่ายถนน และใช้ค่า BC [5] ช่วยในการหาทางแยกที่มีการจราจรคับคั่งสูง และใช้ค่าประสิทธิภาพช่วยในการหาความสามารถในการรองรับปริมาณรถ เพื่อนำมาเปรียบเทียบกันทำให้สามารถวิเคราะห์ลำดับความสำคัญของทางแยกต่างๆ และหาความสัมพันธ์ระหว่างทางแยกกับประสิทธิภาพของทางแยกได้ โดยนำไปประยุกต์ใช้ในอนาคตช่วยให้มีการทำงานอย่างเป็นระบบ และมีประสิทธิภาพมากยิ่งขึ้น

ในงานวิจัยนี้ใช้กรณีศึกษาโครงข่ายถนนภายในจังหวัดนนทบุรี โดยใช้เส้นทางหลักที่ใช้ในการจราจรประกอบด้วยถนนทางหลวงและถนนทางหลวงชนบท ซึ่งจังหวัดนนทบุรีนั้นเป็นจังหวัดที่มีพื้นที่ติดกับกรุงเทพมหานคร จังหวัดพระนครศรีอยุธยา จังหวัดปทุมธานี และจังหวัดนครปฐม จึงทำให้จังหวัดนนทบุรีเป็นทางผ่านในการเชื่อมต่อระหว่างจังหวัดที่อยู่โดยรอบ และถูกใช้เป็นเส้นทางผ่านเข้ากรุงเทพมหานคร ประกอบกับจำนวนประชากรในพื้นที่จังหวัดนนทบุรีมีปริมาณมากทำให้การจราจรบนท้องถนนในบางจุดเกิดการจราจรคับคั่งทั้งที่ถนนภายในจังหวัดนนทบุรีมีจำนวนช่องทางการจราจรที่มีขนาดกว้าง แต่ไม่สามารถแก้ไขปัญหารถติดได้อย่างมีประสิทธิภาพ

ในบทความวิจัยได้แบ่งเนื้อหา ดังนี้ คือ ส่วนที่ 2 ทฤษฎีที่เกี่ยวข้อง ส่วนที่ 3 วิธีดำเนินงาน ส่วนที่ 4 ผลการดำเนินงาน และส่วนที่ 5 สรุปผลและข้อเสนอแนะ

2. ทฤษฎีที่เกี่ยวข้อง

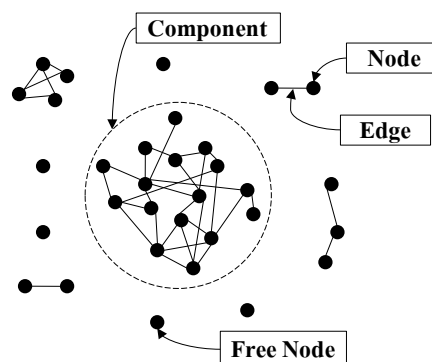
จากการศึกษาการวิเคราะห์โครงข่ายถนนโดยการใช้ทฤษฎีกราฟ มีทฤษฎีที่เกี่ยวข้อง ดังนี้

2.1 ทฤษฎีเครือข่ายซับซ้อน (Complex Networks)

Complex Networks เป็นลักษณะของการเชื่อมต่อเครือข่ายที่มีขนาดใหญ่ และมีความซับซ้อนไม่มีรูปแบบตายตัว โดยเครือข่ายลักษณะนี้มีอยู่ทั่วไปในธรรมชาติ เช่น เครือข่ายของมด [4] เครือข่ายของผึ้ง [4] เครือข่ายทางเศรษฐกิจ หรือแม้กระทั่งเครือข่ายสมองของมนุษย์ โดยมีโครงสร้างหลัก คือ โหนด (Node) และเส้นเชื่อม (Edge) ทำการเชื่อมต่อกัน จึงทำให้เกิดเครือข่ายขนาดใหญ่และซับซ้อนยิ่งขึ้น ซึ่งใน Complex Networks นี้จะมีการเพิ่มและลดของโหนดอยู่ตลอดเวลาทำให้

ไม่สามารถหารูปแบบของเครือข่ายหรือการเชื่อมต่อได้จากงานวิจัย [4] แบ่งรูปแบบ Complex Networks ได้เป็น 3 ประเภท คือ Random Network Model [4], Small-World Network Model [4] และ Scale-Free Network Model [4]

2.1.1 Random Network Model เป็นรูปแบบการเชื่อมต่อของเครือข่ายที่เกิดจากการสุ่มในการเพิ่มโหนดเพื่อเป็นสมาชิกในเครือข่าย โดยขึ้นอยู่กับข้อกำหนดจำนวนโหนด และจำนวนเส้นเชื่อม รูปแบบ Random Network แสดงดังภาพที่ 1

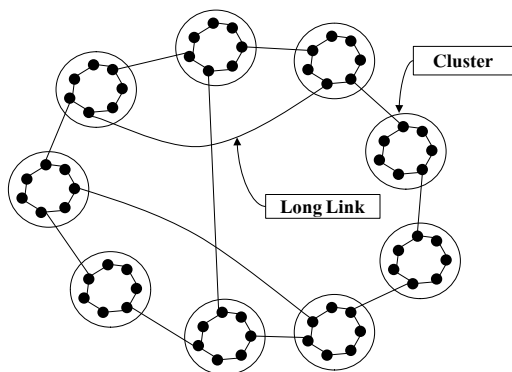


ภาพที่ 1: รูปแบบเครือข่าย Random Network Model

จากภาพที่ 1 ประกอบไปด้วยโหนดที่ไม่ได้เชื่อมต่อกับโหนดอื่นเรียกว่า Free Node และโหนดที่เชื่อมต่อกันเป็นกลุ่มเรียกว่า Component

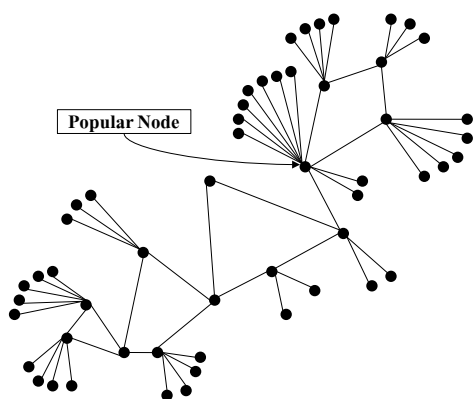
2.1.2 Small-World Network Model เป็นรูปแบบที่ถูกพัฒนามาจาก Random Network Model มีการเชื่อมต่อเป็นเครือข่าย (Cluster) หลายเครือข่าย และในแต่ละเครือข่ายจะเชื่อมโยงกันผ่าน Long Link โดยมีพื้นฐานมาจากเครือข่ายทางสังคมที่ผู้คนที่อยู่ใกล้ชิดกันย่อมมีความสัมพันธ์กันมากกว่าคนที่อยู่ไกลกันออกไป และคนในกลุ่มอาจรู้จักคนอื่นในกลุ่มอื่นที่อยู่ไกลกันออกไป รูปแบบ Small-World Network แสดงดังภาพที่ 2

2.1.3 Scale-Free Network Model เป็นรูปแบบที่มีพื้นฐานมาจากกฎการกระจายตัวแบบเบ้ขวา (Skewed Distribution) ตามกฎแห่งอำนาจ (Power-Law Degree Distribution) โดยโหนดที่มีเส้นเชื่อมต่อกันมาก จะมีอยู่จำนวนน้อย ส่วนโหนดที่มีเส้นเชื่อมต่อน้อย จะมีอยู่จำนวนมาก (Preferential Attachment) ซึ่งการเพิ่มโหนดเพื่อเป็นสมาชิกในเครือข่ายจะทำการเพิ่มกับโหนดที่มีจำนวนเส้นเชื่อมต่อกันมาก (Popular Node) ก่อนเสมอ รูปแบบ Scale-Free Network แสดงดังภาพที่ 3



ภาพที่ 2: รูปแบบเครือข่าย Small-World Network Model

จากภาพที่ 2 ประกอบไปด้วยโหนดที่เชื่อมต่อกันเป็นกลุ่ม เรียกว่า Cluster และเส้นเชื่อมต่อยาวระหว่าง Cluster เรียกว่า Long Link



ภาพที่ 3: รูปแบบเครือข่าย Scale-Free Network Model

จากภาพที่ 2 ประกอบไปด้วยโหนดที่มีจำนวนเส้นเชื่อมต่อมากที่สุดเรียกว่า Popular Node

2.2 ทฤษฎีกราฟ (Graph Theory)

ทฤษฎีกราฟเป็นทฤษฎีที่เกิดขึ้นตั้งแต่ปี ค.ศ. 1736 โดยเริ่มจากปัญหาของสะพานเคอนิกส์เบิร์ก (Königsberg Bridge Problem) ซึ่งเป็นการแก้ปัญหาคณิตศาสตร์โดยอาศัยทฤษฎีกราฟ ใช้โหนด (Vertices) และเส้นเชื่อมต่อยาวระหว่างโหนด (Edge) เชื่อมต่อกัน [3] ซึ่งในการแบ่งประเภทของกราฟจากเส้นเชื่อมต่อยาวระหว่างโหนดสามารถแบ่งได้ 2 ประเภท ได้แก่ กราฟแบบไม่มีทิศทาง (Undirected Graph) ซึ่งมีความสามารถในการส่งข้อมูลไปและกลับได้ในการเชื่อมต่อ และกราฟแบบมีทิศทาง (Directed Graph) ซึ่งมีความสามารถในการส่งหรือรับข้อมูลขึ้นอยู่กับทิศทางที่กำหนดทิศทาง โดยสามารถเขียนนิยามของกราฟ [3] ดังนี้

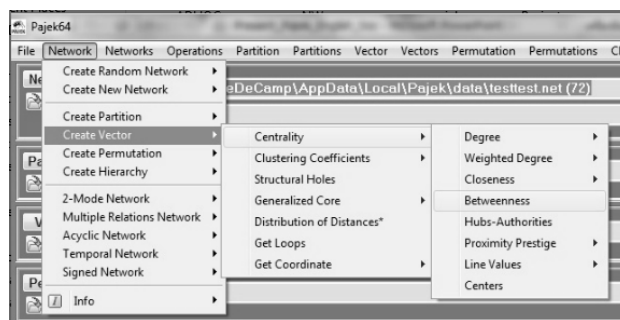
$$G = (V_G, E_G) \tag{1}$$

โดยที่ $E_G \subseteq E_G(V_G)$ ให้ V_G แทนเซตของโหนด ในกราฟ G และให้ E_G แทนเซตของเส้นเชื่อมต่อยาวระหว่างโหนด ในกราฟ G จากกราฟดังกล่าวทำให้สามารถหาค่าคุณลักษณะต่าง ๆ ได้ เช่น ขนาดและความหนาแน่น (Network Size and Network Density) ลักษณะการเชื่อมโยง (Connectivity) หรือค่าความเป็นจุดศูนย์กลาง (Centrality) [5]

2.3 โปรแกรมพาเจค (Pajek)

Pajek เป็นโปรแกรมที่ช่วยในการสร้างแบบจำลอง และวิเคราะห์เครือข่ายขนาดใหญ่ที่มีความซับซ้อนมาก โปรแกรมนี้ได้รับการพัฒนาขึ้นปี ค.ศ. 1996 [6] โดยมีเป้าหมายหลักเพื่อใช้วิเคราะห์ค่าคุณลักษณะ เช่น การหาขนาดและความหนาแน่น ลักษณะการเชื่อมโยง ค่าความเป็นจุดศูนย์กลาง [7] และสร้างแบบจำลองได้อย่างมีประสิทธิภาพ

อีกทั้งมีหลายงานวิจัยที่ใช้โปรแกรม Pajek ในการช่วยจำลองโครงข่ายและวิเคราะห์ข้อมูล Complex Networks ได้แก่ R. Zeng [8] ใช้ในการวิเคราะห์โครงข่ายสังคมออนไลน์ (Social Networks) L. Baoqiang [9] ใช้ในการวิเคราะห์เครือข่ายอินทราเน็ต (Intranet Networks) และ P. Li [10] ทดสอบการใช้งาน Pajek กับ Complex Networks โดยหน้าจอโปรแกรม Pajek แสดงดังภาพที่ 4

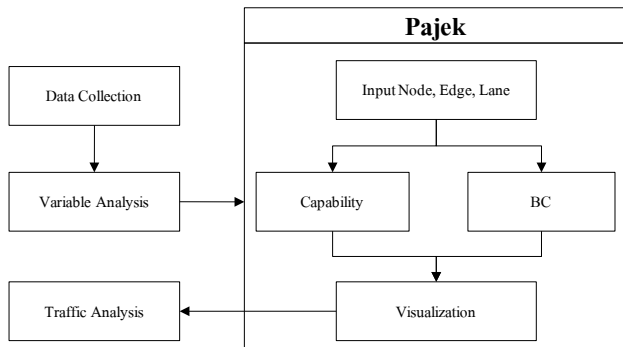


ภาพที่ 4: ตัวอย่างหน้าจอการวิเคราะห์ด้วย Pajek

3. วิธีดำเนินงาน

วิธีดำเนินงานวิจัยสำหรับวิเคราะห์โครงข่ายถนน โดยใช้ค่า Betweenness Centrality (BC) และค่า Capability (C) ได้เตรียมข้อมูลสำหรับการวิเคราะห์คุณลักษณะ การสร้าง

แบบจำลอง และการวิเคราะห์การจราจร โดยมีขั้นตอนการดำเนินงาน แสดงดังภาพที่ 5



ภาพที่ 5: ขั้นตอนการดำเนินงาน

3.1 ข้อมูลที่ใช้ในการดำเนินงาน (Data Collection)

การเก็บข้อมูล โดยใช้ข้อมูลของถนนสายหลักในการเดินทาง ได้แก่ ถนนต่างๆจำนวน 101 เส้น ทางแยกจำนวน 72 ทางแยก และจำนวนช่องทางการจราจรมีค่าในช่วง 1 ถึง 5 ช่องทางการจราจร

3.2 วิเคราะห์คุณลักษณะ

การวิเคราะห์คุณลักษณะใช้ค่า Betweenness Centrality (BC) และค่า Capability (C)

การหาค่า BC จะทำให้หาค่าความเป็นจุดศูนย์กลางโดยวัดจากค่า BC ที่หามาได้ คือ โหนดที่มีความสำคัญที่ต้องทำการผ่านบ่อยเพื่อไปยังจุดอื่นภายในเครือข่าย [6] โดยมีสมการดังนี้

$$BC_u = \sum \frac{\sum_{i \in S_{ij}} \delta_i^u}{|S_{ij}|} \quad (2)$$

โดยให้ $\in S_{ij} \delta_i^u$ แทนจำนวนระยะทางที่สั้นที่สุดจาก i ไป j ที่ผ่านจุด δ_i^u และให้ S_{ij} แทนจำนวนระยะทางที่สั้นที่สุดจาก i ไป j

การหาค่า C วิเคราะห์จากข้อมูลของทางแยก เช่น จำนวนทางแยก จำนวนช่องทางการจราจร และระยะทางระหว่างแยก นำมาหาความสัมพันธ์และสร้างสมการใช้ในการคำนวณค่าประสิทธิภาพ โดยมีสมการดังนี้

$$C_i = \frac{\sum L(E_i)}{\sum E_i} \quad (3)$$

โดยให้ $L(E_i)$ แทนด้วยจำนวนช่องทางจราจรในแต่ละถนนของแยก i และให้ E_i แทนด้วยจำนวนถนนของแยก i

3.3 การสร้างแบบจำลองโดยใช้โปรแกรม Pajek

ในขั้นตอนนี้ทำการแปลงข้อมูลให้ในรูปแบบที่โปรแกรม Pajek สามารถอ่านค่าได้ ซึ่งทำการกำหนดให้ทางแยกเป็นโหนด และกำหนดให้ถนนที่เชื่อมต่อระหว่างทางแยกเป็นเส้นเชื่อม โดยให้กราฟเป็นแบบกราฟไม่มีทิศทาง (Undirected Graph) เนื่องจากถนนมีช่องทางการจราจร 2 ช่องทางสามารถเดินทางไปกลับได้

3.4 การวิเคราะห์การจราจร (Traffic Analysis)

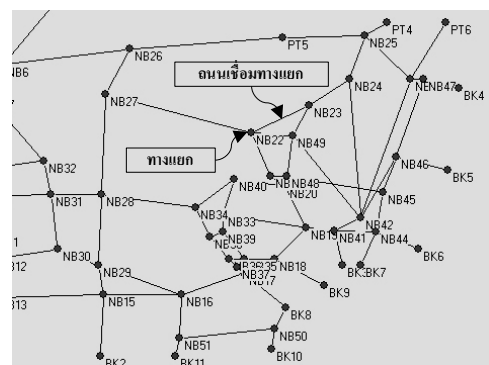
ในขั้นตอนของการวิเคราะห์ข้อมูลที่ได้จากแบบจำลองแบ่งการวิเคราะห์ออกเป็น 2 ส่วน คือ การวิเคราะห์หาค่า BC โดยใช้สมการที่ (2) และค่า C โดยใช้สมการที่ (3) เมื่อได้ค่า BC และค่า C แล้วนำค่าทั้งสองมาเปรียบเทียบหาความสัมพันธ์

4. ผลการดำเนินงาน

คณะผู้วิจัยได้ทำการจำลองโครงข่ายถนน วิเคราะห์คุณลักษณะ และผลการเปรียบเทียบค่า BC กับค่า C

4.1 ผลการจำลองโครงข่ายถนน

ในการวิเคราะห์โครงข่ายถนน โดยใช้ค่า BC และค่า C ผลการสร้างแบบจำลองโครงข่ายโดยใช้โปรแกรม Pajek สามารถทำการสร้างโครงข่ายถนนจำลองได้อย่างถูกต้องตามข้อมูลถนนจริง แสดงดังภาพที่ 6



ภาพที่ 6: ตัวอย่างการจำลองโครงข่ายถนน

โดยใช้จุดวงกลมแสดงทางแยกต่างๆ เส้นตรงเชื่อมต่อระหว่างจุดแสดงถนน และใช้การกำหนดชื่อทางแยกด้วยรหัสซึ่งจะได้

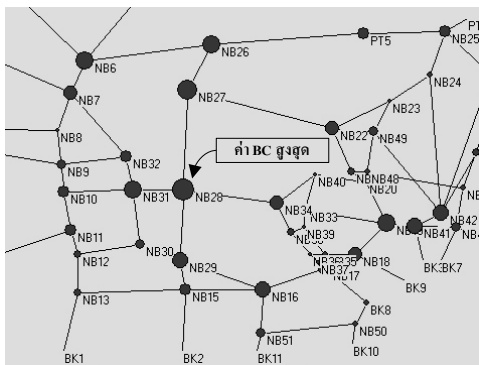
$$G_{street} = (V_{street}, E_{street})$$

โดย แทนทางแยก และ แทนถนนเชื่อมระหว่างทาง

แยก ซึ่งมี = 72 และ = 101

4.2 ผลการวิเคราะห์คุณลักษณะ

ตัวอย่างผลการจำลองโครงข่ายโดยใช้ค่า BC ร่วมกับโครงข่ายถนน โดยสามารถแสดงจากค่า BC ที่มีค่ามาก จุดวงกลมจะมีขนาดใหญ่ แต่ถ้าค่า BC ที่มีค่าน้อย จุดวงกลมจะมีขนาดเล็ก โดยแสดงดังภาพที่ 7



ภาพที่ 7: ตัวอย่างค่า BC ของโครงข่ายถนน

โดยได้ค่าเฉลี่ยของค่า BC เท่ากับ 0.258 สามารถเรียงลำดับค่า BC จากมากไปน้อย 5 อันดับแรก ดังแสดงในตารางที่ 1

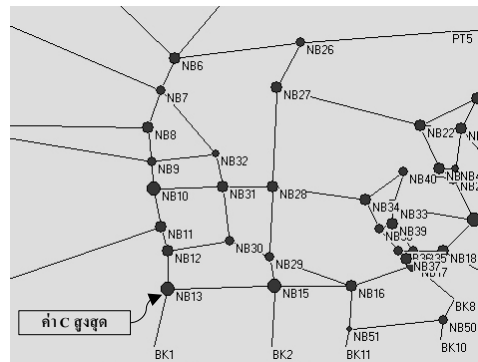
ตารางที่ 1: ทางแยกที่มีค่า BC มากที่สุด 5 อันดับแรก

No	Vertices	Crossroad	BC
1	NB28	แยกถนนราชพฤกษ์ตัดถนนรัตนวิเบศร์	0.318
2	NB27	แยกถนนราชพฤกษ์ตัดถนนชัยพฤกษ์	0.245
3	NB26	แยกถนนราชพฤกษ์ตัดถนนบางบัวทอง	0.218
4	NB31	แยกถนนรัตนวิเบศร์ตัดถนนบางกรวย-ไทรน้อย	0.205
5	NB19	แยกแคราย	0.187

จากตารางที่ 1 แสดงการเรียงลำดับของค่า BC ซึ่งมีค่ามากแสดงถึงเป็นทางแยกที่มีรถผ่านมาก โดยค่าที่มากที่สุด คือ จุด

NB28 ทางแยกถนนราชพฤกษ์ตัดถนนรัตนวิเบศร์ โดยมีค่าเท่ากับ 0.318

ตัวอย่างผลการจำลองโครงข่ายโดยใช้ค่า C ร่วมกับโครงข่ายถนน โดยสามารถแสดงค่าประสิทธิภาพที่มีค่ามาก จุดวงกลมจะมีขนาดใหญ่ แต่ถ้าค่าประสิทธิภาพที่มีค่าน้อย จุดวงกลมจะมีขนาดเล็ก โดยแสดงดังภาพที่ 8



ภาพที่ 8: ตัวอย่างค่าประสิทธิภาพของโครงข่ายถนน

โดยค่าเฉลี่ยของค่า C เท่ากับ 0.447 สามารถเรียงลำดับค่า C จากมากไปน้อย 5 อันดับแรก ดังแสดงในตารางที่ 2

ตารางที่ 2: ทางแยกที่มีค่าประสิทธิภาพมากที่สุด 5 อันดับแรก

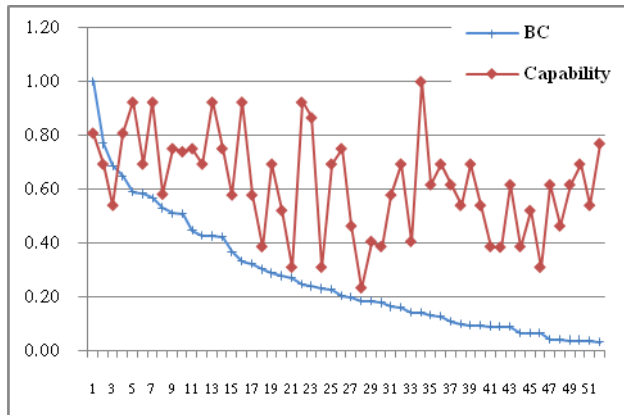
No	Vertices	Crossroad	C
1	NB13	แยกถนนกาญจนาตัดถนนนครอินทร์	1.000
2	NB10	แยกถนนกาญจนาตัดถนนรัตนวิเบศร์	0.923
3	NB15	แยกถนนนครอินทร์ตัดถนนราชพฤกษ์	0.923
4	NB17	แยกพระราม5	0.923
5	NB19	แยกแคราย	0.923

จากตารางที่ 2 แสดงการเรียงลำดับของค่าประสิทธิภาพ ซึ่งมีค่ามากที่สุด แสดงถึงมีความสามารถในการระบายรถหรือรองรับปริมาณรถได้มาก โดยค่าที่มากที่สุด คือ จุด NB13 ทางแยกถนนกาญจนาตัดถนนนครอินทร์ โดยมีค่าเท่ากับ 1.000

4.3 ผลการเปรียบเทียบค่า BC กับค่า C

เมื่อนำทั้งสองค่ามาเปรียบเทียบกันแล้ว พิจารณาถึงค่า BC ในตารางที่ 1 มีค่า C ดังนี้ จุด NB28 (แยกถนนราชพฤกษ์ตัดถนนรัตนวิเบศร์) ค่า $BC_{NB28} = 0.318$ ค่า $C_{NB28} = 0.807$ จุด NB27 (แยกถนนราชพฤกษ์ตัดถนนชัยพฤกษ์) ค่า $BC_{NB27} = 0.245$ ค่า $C_{NB27} = 0.692$ จุด NB26 (แยกถนนราชพฤกษ์ตัดถนนบางบัวทอง) ค่า $BC_{NB26} = 0.218$ ค่า $C_{NB26} = 0.538$ จุด NB31

(แยกถนนริเบตตัดถนนบางกรวย-ไทรน้อย) ค่า $BC_{NB31} = 0.205$ ค่า $C_{NB31} = 0.807$ และจุด NB19 (แยกแคราย) ค่า $BC_{NB19} = 0.187$ ค่า $C_{NB19} = 0.923$ ซึ่งเห็นได้ว่าค่า C ของทางแยกดังกล่าวไม่อยู่ใน 5 อันดับแรกทั้งหมด เนื่องจากค่า C มากในทางแยกที่มีค่า BC น้อย โดยแสดงดังภาพที่ 9



ภาพที่ 9: กราฟเปรียบเทียบระหว่าง BC กับ Capability

จากภาพที่ 9 นำเสนอได้ว่าทางแยกใดมีค่า BC สูงและมีค่า C สูง ส่งผลให้มีความสามารถในระบายนรถได้ดี ถ้าทางแยกใดมีค่า BC สูงและมีค่า C ต่ำ ส่งผลให้เป็นสาเหตุหนึ่งที่ทำให้เกิดปัญหาการติดในจุดดังกล่าว และถ้าทางแยกใดมีค่า BC ต่ำและมีค่า C ต่ำ ส่งผลให้มีความสามารถในระบายนรถได้ดี

5. สรุปผลและข้อเสนอแนะ

การวิเคราะห์โครงข่ายถนน ที่ได้นำเอาทฤษฎีกราฟและทฤษฎี Complex Networks มาช่วยในการสร้างแบบจำลอง และวิเคราะห์คุณลักษณะ ทำให้ทราบอัตราการไหลเวียนของรถในแยกต่าง ๆ ภายในจังหวัดนนทบุรี พบว่าถ้าทางแยกที่มีความสำคัญแต่มีค่าประสิทธิภาพน้อยจะทำให้เกิดปัญหาการติดได้

เนื่องจากค่า C พิจารณาจากช่องทางจราจรเพียงอย่างเดียว ซึ่งหากมีการเพิ่มตัวแปรในการคำนวณหาค่า C ในการวิเคราะห์ เช่น สัญญาณไฟจราจร ทางลาดชัน อุโมงค์หรือสะพานข้ามทางแยก เพื่อให้ค่า C บอกลถึงประสิทธิภาพของทางแยกได้ดียิ่งขึ้น และควรทำการทดลองในกรณีที่ทางแยกใดทางแยกหนึ่งเกิดปัญหาไม่สามารถใช้งานได้ แล้วภาวะของการเป็นค่า BC มีการย้ายไปอยู่ที่ไหนใด เพื่อให้ง่ายในการจัดการวางแผนรับมือกับจำนวนรถ

เอกสารอ้างอิง

- [1] ชูกร จุลินทร, “ความเชื่อมโยงระหว่างกันในอาเซียน (ASEAN Connectivity) : กรณีความเชื่อมโยงทางกายภาพของภาคคมนาคมทางบก,” *เอกสารข่าวสารงานวิจัยและพัฒนา* ปีที่ 11 ฉบับที่ 121 พ.ศ. 2555 หน้า 22–56.
- [2] B. Lucas, “Which side of the road do they drive on?” <http://www.brianlucas.ca/roadside>. 20-Aug-2013.
- [3] T. Harju, *GRAPH THEORY*, Department of Mathematics, University of Turku: Finland, 2012.
- [4] W. Xiao Fan and C. Guanrong, “Complex Networks : Small-World, Scale-Free and Beyond,” *IEEE Circuits and Systems Magazine*, vol. 2003, pp. 6–20.
- [5] รุจเรชา วิทยายุทธินกุล และ น้าทิพย์ วิภาวิน, “การวิเคราะห์เครือข่ายทางสังคม กับงานวิจัยทางสารสนเทศศาสตร์,” *วารสารบรรณศาสตร์ มศว* ปีที่ 5 ฉบับที่ 2 พ.ศ. 2555 หน้า 126–140.
- [6] X. Tan, T. Tian, and G. Liu, “Features analysis of small-cell lung cancer signal transduction network by method of complex network,” *International Conference on Future*, vol. 2009, pp. 9–12.
- [7] Vladimir Batagelj and Andrej Mrvar, *Pajek Program for Analysis and Visualization of Large Networks : Reference Manual*, Ljubljana, 2011.
- [8] R. Zeng, H. Shen, and T. W. Xu, “A simple model to characterize social networks,” *18th IEEE International Conference on Networks (ICON)*, 2012, pp. 13–17.
- [9] L. Baoqiang, T. Shurong, S. Shoukui, and M. Cuijing, “On the simulation of the network topology generator and robustness of the constructed network,” *2nd International Conference on Intelligent Control and Information Processing (ICICIP)*, 2011, vol. 2, pp. 725–728.
- [10] P. Li, H. Zhao, Y. Qiao, Z. Liu, H. Li, and B. Li, “A platform of software network measurement design and implement,” *2nd International Conference on Advanced Computer Control (ICACC)*, 2010, vol. 2, pp. 336–339.

การพัฒนาระบบตรวจแบบฝึกหัดการเขียนโปรแกรมแบบอัตโนมัติ

The Development of Automatic Programming Exercise Verification System

ศราวุธ รุ่งเจริญกิจ (Sarawuth Rungcharoenkit)¹ และ กิติ์สุชาติ พสุภา (Kitsuchart Pasupa)²

^{1,2}คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
iam@balll.me¹, kitsuchart@it.kmitl.ac.th²

บทคัดย่อ

บทความนี้นำเสนอการพัฒนาระบบตรวจแบบฝึกหัดการเขียนโปรแกรมแบบอัตโนมัติ ซึ่งเป็นเครื่องมือสำหรับตรวจสอบข้อมูลส่งออกของโปรแกรมจากชุดข้อมูลนำเข้าที่กำหนดโดยอัตโนมัติ ช่วยอำนวยความสะดวกในการตรวจคำตอบของแบบฝึกหัดในการเรียนวิชาเขียนโปรแกรม มีวัตถุประสงค์เพื่อให้ผู้เรียนสามารถตรวจคำตอบได้อย่างรวดเร็วด้วยตนเอง ส่งผลให้มีความต่อเนื่องในการทำแบบฝึกหัดและปรับปรุงโปรแกรมของตนเอง ปัจจุบันการตรวจแบบฝึกหัดการเขียนโปรแกรมในชั้นเรียนจะเป็นการส่งรหัสต้นฉบับให้ผู้สอนตรวจหรือยกมือรอมผู้สอนมาตรวจที่คอมพิวเตอร์ของตน ซึ่งการรอตรวจแต่ละครั้งใช้ระยะเวลาอันยาวนานและมีความแม่นยำต่ำเนื่องจากผู้สอนมีจำนวนน้อยกว่าผู้เรียน ดังนั้นการพัฒนาระบบนี้จะช่วยให้ผู้เรียนสามารถตรวจคำตอบของแบบฝึกหัดการเขียนโปรแกรมของตนเองได้รวดเร็วขึ้น โดยระบบประกอบด้วย 5 ส่วน คือ (ก) เว็บแอปพลิเคชัน (ข) ระบบการตรวจโปรแกรม (ค) ระบบแซนด์บ็อกซ์ (ง) ระบบตรวจสอบการคัดลอก (จ) ระบบการสอบ โดยระบบทั้งหมดถูกพัฒนาด้วยภาษาไพทอน

คำสำคัญ: การตรวจแบบฝึกหัด การเขียนโปรแกรม เครื่องมือการสอน แซนด์บ็อกซ์

Abstract

This paper presents the development of automatic programming exercise verification system which can automatically compare submitted program's output with a predefined answer and specific input. The objective is to allow students to quickly check their answer themselves. So they can continuously do their exercises. Currently, students must wait for the instructors to check

their programs on their computers. This action is time consuming and has less accuracy because the number of instructors is always less than the number of students. So the development of automatic programming exercise verification system will allow students to check their program immediately after they have submitted. The system is developed by using python language and consists of five parts: (i) Web application (ii) Grader (iii) Sandbox (iv) Copy paste detector, and (v) Examination system.

Keyword: Exercise Marking Programming Teaching Tool Sandbox.

1. บทนำ

เทคโนโลยีได้เข้ามามีบทบาทกับชีวิตประจำวันมากขึ้น การศึกษาการพัฒนาซอฟต์แวร์จึงเป็นสิ่งสำคัญ ในปัจจุบัน การศึกษาด้านการพัฒนาซอฟต์แวร์ในห้องเรียนจะเน้นการปฏิบัติโดยการให้โจทย์แบบฝึกหัดแก่ผู้เรียนทดลองทำ ซึ่งการส่งแบบฝึกหัดนั้นผู้เรียนอาจยกมือส่งสัญญาณให้ผู้สอนเข้ามาตรวจโดยการใส่ข้อมูลนำเข้าหลายประเภทลงไปเพื่อทดสอบความรัดกุมของโปรแกรมที่ผู้เรียนพัฒนา แต่เนื่องจากจำนวนผู้สอนมีน้อยกว่าผู้เรียนมากและการตรวจแต่ละครั้งต้องใช้เวลา ทำให้ไม่สามารถตรวจแบบฝึกหัดของผู้เรียนทุกคนได้อย่างละเอียดเท่ากันในเวลาจำกัด อีกทั้งเมื่อผู้เรียนทราบผลว่าโปรแกรมของตนเองไม่สามารถให้คำตอบที่ถูกต้องในชุดข้อมูลนำเข้าที่นำมาทดสอบได้ทุกชุด ผู้เรียนจะขาดความต่อเนื่องในการพัฒนาโปรแกรมเนื่องจากต้องเสียเวลารอการตรวจทำให้เกิดความล่าช้าและใช้เวลาในคาบเรียนไม่คุ้มค่า ในกรณีที่ผู้สอนมอบหมายแบบฝึกหัดให้เป็นการบ้าน ผู้เรียนจะไม่

สามารถทราบได้เลยว่าโปรแกรมที่ตนเองพัฒนานั้นสามารถให้คำตอบได้ถูกต้องตามชุดข้อมูลนำเข้าทดสอบที่ผู้สอนออกแบบไว้ได้หรือไม่จนกระทั่งคาบเรียนถัดไป ทำให้เกิดความไม่ต่อเนื่องในการเรียนรู้และพัฒนา ซึ่งในปัจจุบันได้มีโครงการ Café-Grader [1] ที่พัฒนาระบบตรวจแบบฝึกหัดการเขียนโปรแกรมแบบอัตโนมัติขึ้นมาใช้งาน แต่ยังคงขาดความสมบูรณ์ในความสามารถหลายด้าน เช่น ระบบการบริหารจัดการโจทย์ ระบบบริหารจัดการบัญชีผู้ใช้ ระบบรายงานคะแนน ระบบตรวจสอบความคล้ายคลึงของโปรแกรมที่ผู้เรียนส่ง ระบบการสอบและยังสามารถตรวจโปรแกรมได้เพียงภาษาเดียว โปรแกรม บทความนี้จึงนำเสนอการพัฒนาแบบฝึกหัดการเขียนโปรแกรมแบบอัตโนมัติ ดำเนินการพัฒนาโดยใช้เทคโนโลยีเว็บเพื่อให้สามารถเข้าถึงได้จากทุกที่ทั้งในห้องเรียนและนอกห้องเรียน โดยมีเป้าหมายเพื่อให้ผู้เรียนสามารถตรวจโปรแกรมของตนเองได้อย่างรวดเร็วและให้ผู้สอนสามารถบริหารจัดการ โจทย์รวมถึงเรียกดูรายงานคะแนนของผู้เรียนได้อย่างง่ายดาย

2. ทฤษฎีและงานที่เกี่ยวข้อง

ระบบตรวจแบบฝึกหัดการเขียนโปรแกรมแบบอัตโนมัติที่นำเสนอมีจุดเด่นที่ ระบบตรวจสอบการคัดลอก และระบบรักษาความปลอดภัยจากการทำงานของโปรแกรมภายนอก (ระบบแซนด์บ็อกซ์) ซึ่งมีรายละเอียดดังนี้

2.1 ระบบตรวจสอบการคัดลอก

ระบบตรวจสอบการคัดลอกของระบบตรวจแบบฝึกหัดการเขียนโปรแกรมแบบอัตโนมัติแบ่งได้ออกเป็น 2 แบบคือ

2.1.1 การตรวจสอบค่าแฮชของไฟล์ สามารถทำได้โดยใช้ อัลกอริทึมคำนวณที่สามารถคำนวณค่าที่มีความยาวใด ๆ ได้ ออกมาเป็นข้อมูลขนาดคงที่ เช่น การขึ้นตอนวิธีคำนวณแบบ MD5 [2] หรือ SHA-1 [3] โดยข้อมูลนำเข้าที่เหมือนกันจะคำนวณได้ค่าแฮชที่เหมือนกันเสมอ สามารถนำมาใช้ประโยชน์ได้ในการตรวจสอบความเหมือนกันของไฟล์ที่ผู้ใช้นำส่งขึ้นมา

2.1.2 Copy Paste Detector (CPD) [4] เป็นซอฟต์แวร์ส่วนหนึ่งในชุด PMD [5] เป็นซอฟต์แวร์วิเคราะห์รหัสต้นฉบับที่แจกจ่ายแบบเปิดเผยรหัสต้นฉบับ CPD สามารถตรวจสอบความคล้ายคลึงของรหัสต้นฉบับ โปรแกรมได้โดยกำหนดค่า

ระดับความคล้ายคลึง (Threshold) เพื่อให้ CPD ดำเนินการวิเคราะห์รหัสต้นฉบับของโปรแกรมที่เป็นข้อมูลนำเข้าว่ามีรหัสชุดใดมีความคล้ายคลึงมากกว่าระดับที่กำหนดบ้าง โดยจะรายงานออกมาเป็นคู่ของชื่อไฟล์ที่คล้ายคลึงกัน เลขที่บรรทัดและรหัสต้นฉบับที่คล้ายคลึงดังแสดงในภาพที่ 1

File	Line
net/sourceforge/pmd/lang/java/ast/JavaParser.java	6457
net/sourceforge/pmd/lang/plsql/ast/PLSQLParser.java	31949

```

}

private boolean jj_2_1(int xla) {
    jj_la = xla; jj_lastpos = jj_scanpos = token;
    try { return !jj_3_1(); }
    catch(LookaheadSuccess ls) { return true; }
    finally { jj_save(0, xla); }
}

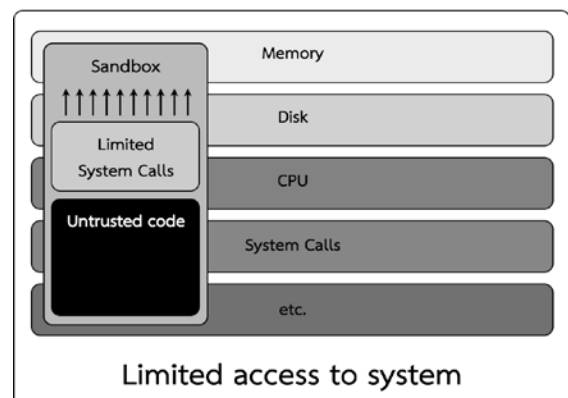
private boolean jj_2_2(int xla) {
    jj_la = xla; jj_lastpos = jj_scanpos = token;
    try { return !jj_3_2(); }
    catch(LookaheadSuccess ls) { return true; }
    finally { jj_save(1, xla); }
}

```

ภาพที่ 1: แสดงผลการทำงานของ CPD

2.2 ระบบแซนด์บ็อกซ์

ระบบแซนด์บ็อกซ์ (Sandbox) คือกลไกทางด้านความปลอดภัยแบบหนึ่งของระบบคอมพิวเตอร์ ใช้เพื่อแบ่งแยกโปรแกรมที่กำลังทำงานออกไปในอีกสภาพแวดล้อมที่ถูกควบคุมอย่างจำกัดเพื่อป้องกันไม่ให้โปรแกรมที่ไม่น่าเชื่อถือสามารถทำอันตรายกับระบบหรือข้อมูลอื่น ๆ ได้ โดยการจำกัดการทำงาน การเข้าถึงทรัพยากรและการเข้าถึงข้อมูลของโปรแกรมที่อยู่ในแซนด์บ็อกซ์ เพื่อไม่ให้เกิดผลกระทบต่อทรัพยากรโดยรวมในกรณีที่โปรแกรมในแซนด์บ็อกซ์ทำอันตรายต่อระบบดังแสดงในภาพที่ 2



ภาพที่ 2: แสดงการทำงานของระบบแซนด์บ็อกซ์

3. แนวคิดและการดำเนินงาน

ระบบตรวจแบบฝึกหัดการเขียน โปรแกรมแบบอัตโนมัติ ออกแบบโดยแบ่งออกเป็น 5 ส่วน คือ เว็บแอปพลิเคชัน (Web Application) ระบบการตรวจ (Grader) ระบบแซนด์บ็อกซ์ (Sandbox) ระบบตรวจสอบการคัดลอก (Copy Paste Detector) และ ระบบการสอบ (Exam System) โดยระบบทั้งหมดพัฒนาด้วยภาษาไพทอนดังภาพที่ 3 และมีรายละเอียดแต่ละระบบดังต่อไปนี้

3.1 เว็บแอปพลิเคชัน

เว็บแอปพลิเคชันเป็นส่วนติดต่อผู้ใช้สำหรับผู้สอนและผู้เรียน (ภาพที่ 4) มีการแบ่งระดับของผู้ใช้เป็น 2 ระดับคือ ผู้สอน (Instructor) และ ผู้เรียน (Student) โดยในระบบประกอบด้วยฟังก์ชันการทำงานดังภาพที่ 5

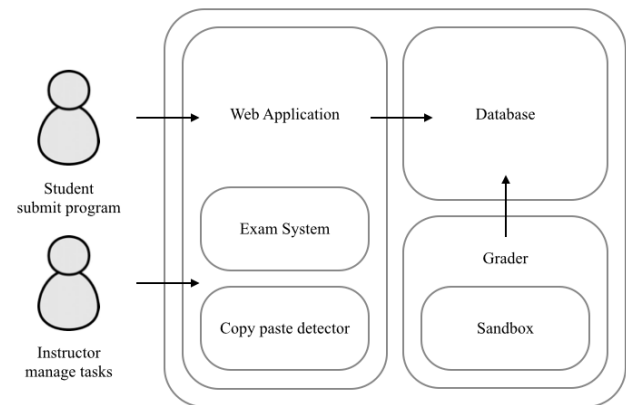
3.1.1 ผู้สอนสามารถบริหารจัดการโจทย์แบบฝึกหัด บริหารจัดการชุดข้อมูลนำเข้าทดสอบและคำตอบ เรียกดูรายงานคะแนน รายงานการตรวจสอบการคัดลอกของผู้เรียนและเรียกดูรหัสต้นฉบับของโปรแกรมที่ผู้เรียนนำส่งเข้าสู่ระบบ

3.1.2 ผู้เรียนสามารถนำส่งโปรแกรมเข้าสู่ระบบเพื่อขอรับการตรวจ เรียกดูรายการการส่งตรวจทั้งหมด เรียกดูรายละเอียดและผลลัพธ์การตรวจ เรียกดูรายการโจทย์แบบฝึกหัดทั้งหมด (ภาพที่ 6) เรียกดูรายละเอียดของโจทย์แบบฝึกหัด (ภาพที่ 7) และเรียกดูการจัดอันดับคะแนนของผู้เรียนทั้งหมดได้ ซึ่งผู้สอนจะสามารถใช้งานฟังก์ชันของผู้เรียนได้เหมือนกัน

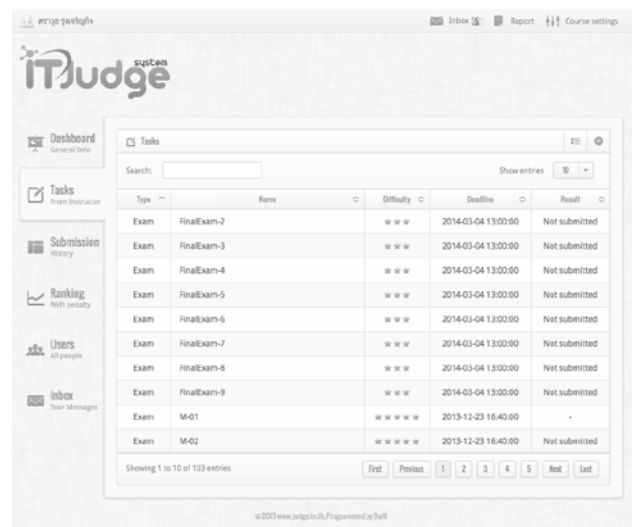
3.2 ระบบการตรวจ

ระบบสามารถตรวจโปรแกรมที่ผู้เรียนนำส่งเข้ามาจากเว็บแอปพลิเคชัน โดยจะทำงานผ่านส่วนติดต่อผู้ใช้แบบคอมมานด์ไลน์ซึ่งจะทำการคอมไพล์โปรแกรกดังกล่าวในระบบแซนด์บ็อกซ์ โดยสามารถกำหนดเงื่อนไขการคอมไพล์ได้ เช่น จะต้องมีการกำหนดในรหัสต้นฉบับหรือจะต้องไม่มีคำที่กำหนดในรหัสต้นฉบับ จึงจะอนุญาตให้ดำเนินการคอมไพล์ได้ เพื่อเป็นการบังคับให้ผู้เรียนเขียนโปรแกรมตามที่กำหนด เช่น ห้ามใช้คำสั่ง if เป็นต้น จากนั้นระบบจะนำชุดข้อมูลนำเข้าทดสอบ (Test case) จากฐานข้อมูลเป็นข้อมูลนำเข้า (Input) ของโปรแกรม เมื่อระบบการตรวจได้รับข้อมูลส่งออก (Output) จากการทำงานของโปรแกรมจะนำไปเปรียบเทียบกับเฉลย (Answer) ข้อมูลส่งออกของโจทย์ข้อนั้นในฐานข้อมูลโดยไม่

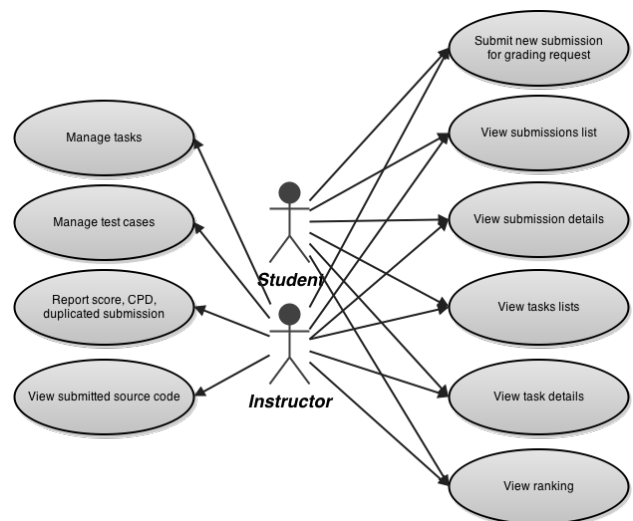
สนใจช่องว่างทางขวาของแต่ละบรรทัด จากนั้นจึงคำนวณเป็นคะแนนและเวลาที่โปรแกรมใช้ทำงานดังภาพที่ 8



ภาพที่ 3: ภาพรวมของระบบ



ภาพที่ 4: ส่วนติดต่อผู้ใช้งานเว็บแอปพลิเคชัน



ภาพที่ 5: Use Case Diagram ของระบบ

จากนั้นจะนำไปจัดอันดับและบันทึกผลการตรวจลงฐานข้อมูลต่อไป การจัดอันดับของผู้เรียนในระบบนั้นสามารถจัดอันดับได้จากคะแนนรวมของผู้เรียนในระบบ (*TotalScore*) โดยคะแนนรวมสามารถคำนวณได้จาก

$$TotalScore = nCorrect \times 10 \quad (1)$$

เมื่อ *nCorrect* คือจำนวนชุดทดสอบที่คำตอบถูกต้อง

หากผู้เรียนมีคะแนนเท่ากันจะพิจารณาเพิ่มเติมจากค่าเหล่านี้ตามลำดับโดยพิจารณาจากค่าเหล่านี้ตามลำดับ

- จำนวนโจทย์แบบฝึกหัดที่ทำถูกครบทุกชุดทดสอบ (*TotalPassedTask*)

- ระดับความยากรวมของโจทย์ที่มีคำตอบถูกครบทุกชุดทดสอบ (*TotalStar*) โดยสามารถคำนวณได้จาก

$$TotalStar = \sum PassedStar_i \quad (2)$$

เมื่อ *PassedStar_i* คือระดับความยากของโจทย์ที่มีคำตอบถูกต้องในทุกชุดทดสอบ

- ค่าปรับรวม (*TotalPenalty*) สามารถคำนวณได้จาก

$$TotalPenalty = \sum TaskPenalty_i \quad (3)$$

เมื่อ *TaskPenalty_i* คือค่าปรับของโจทย์แต่ละข้อที่ผู้ใช้นำส่งซึ่งสามารถคำนวณได้จาก

$$TaskPenalty = t_{Create} - t_{LastSubmissionTime} \quad (4)$$

เมื่อ t_{Create} คือเวลาที่สร้างโจทย์ และ $t_{LastSubmissionTime}$ คือเวลาที่ทำการส่ง โดยเวลาที่นำมาคำนวณเป็นเวลาแบบ Unix POSIX Time หน่วยเป็นวินาที

สำหรับภาษาการเขียนโปรแกรมที่รองรับจะขึ้นอยู่กับกระบวนการทำงานของภาษาและการรองรับการจำกัดการทำงานของระบบแซนด์บ็อกซ์ ซึ่งในปัจจุบันระบบสามารถรองรับการเขียนโปรแกรมได้ 3 ภาษาคือ ภาษาไพทอน ภาษาซีและภาษาเรปเตอร์ [6]

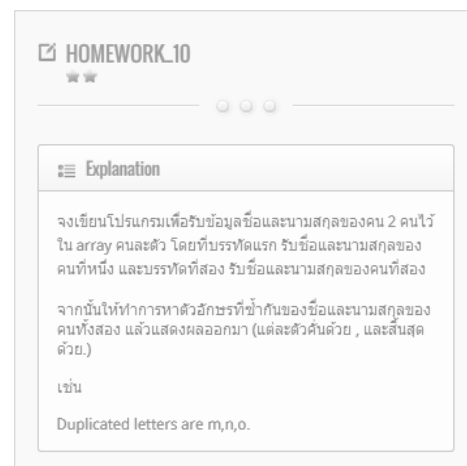
3.3 ระบบแซนด์บ็อกซ์

เพื่อไม่ให้โปรแกรมของผู้เรียนสามารถใช้งานทรัพยากรของเครื่องจนหมดหรือทำอันตรายต่อระบบได้จึงมีการจำกัดการเวลาการเข้าใช้งานหน่วยประมวลผลกลาง ขนาดของหน่วยความจำหลัก สิทธิในการอ่านเขียนไฟล์ในระบบและสิทธิในการเรียกใช้งานคำสั่งเรียกระบบ การตรวจแต่ละครั้งสภาพแวดล้อมจำลองของระบบแซนด์บ็อกซ์จะถูกทำลายและ

สร้างขึ้นใหม่เสมอเพื่อให้การตรวจแต่ละครั้งมีมาตรฐานการตรวจเท่ากันและทำให้ผู้ใช้ไม่สามารถเก็บไฟล์ไว้ในสภาพแวดล้อมจำลองเพื่อให้ผู้ใช้ย้อนนำไปใช้งานต่อได้อีกด้วย

Type	Name	Difficulty
Exam	FinalExam-2	***
Exam	FinalExam-3	***
Exam	FinalExam-4	***
Exam	FinalExam-5	***
Exam	FinalExam-6	***
Exam	FinalExam-7	***
Exam	FinalExam-8	***
Exam	FinalExam-9	***
Exam	M-01	*****
Exam	M-02	*****

ภาพที่ 6: ส่วนติดต่อผู้ใช้แสดงรายการ โจทย์แบบฝึกหัด



ภาพที่ 7: ส่วนติดต่อผู้ใช้แสดงรายละเอียด โจทย์แบบฝึกหัด

Submission#37643	
Submitter:	นายณัฐกิตติ์ วีชรเชียรสกุล (it56070036)
Language:	C
Task:	FinalExam-7
Result:	PP-
Points:	20
Execution time:	0.00736999511719 seconds
Submission time:	2014-03-04 12:57:31

ภาพที่ 8: แสดงข้อมูลที่ได้รับการตรวจ

3.4 ระบบตรวจสอบความคล้ายคลึงของโปรแกรม

เพื่อค้นหาความคล้ายคลึงของโปรแกรมที่ผู้เรียนนำส่งเข้ามาจึงมีระบบการตรวจสอบความคล้ายคลึงของโปรแกรม ประกอบด้วย การตรวจสอบ 2 ขั้นตอนคือ การตรวจสอบค่าแฮชของไฟล์ และการใช้งานระบบ CPD ในการตรวจสอบ เมื่อผู้นำส่งโปรแกรมขึ้นมาบนระบบจะถูกตรวจสอบค่าแฮชโดยนำไปเปรียบเทียบกับค่าแฮชของการส่งครั้งก่อน ๆ ในฐานข้อมูล หากพบค่าแฮชซ้ำจะทำการแจ้งเตือนให้ผู้สอนทราบผ่านทางระบบข้อความส่วนตัวและผู้สอนสามารถเรียกดูสรุปการตรวจสอบค่าแฮชได้เองผ่านการเรียกดูรายงานของระบบ ดังภาพที่ 9 สำหรับการตรวจสอบความคล้ายคลึงด้วยระบบ CPD นั้นผู้สอนสามารถดำเนินการตรวจได้ผ่านทาง การเรียกดูรายงานของระบบเช่นกันดังแสดงในภาพที่ 10

Hash	Task	SubID	User
5976cf0ebb7ef4f101da9d77cc2cae385a5f2b9	Lab 1-1	13483	it56070018
5976cf0ebb7ef4f101da9d77cc2cae385a5f2b9	Lab 1-1	13615	it56070030

ภาพที่ 9: แสดงผลลัพธ์การตรวจการคัดลอกด้วยค่าแฮช

```
Starting at line 2 of 15481.c
Starting at line 2 of 15464.c
Starting at line 2 of 13980.c

int main()
{
printf("Hello");
printf("IT KMITL");
return 0;
}
=====
Found a 4 line (16 tokens) duplication in the following files:
Starting at line 2 of 13863.c
Starting at line 2 of 13972.c

int main(){
printf("Hello");
printf("IT KMITL");
}
=====
Found a 4 line (15 tokens) duplication in the following files:
Starting at line 2 of 14470.c
Starting at line 2 of 15361.c
Starting at line 2 of 13818.c
Starting at line 2 of 14754.c
Starting at line 2 of 15061.c
Starting at line 2 of 14789.c
Starting at line 2 of 15429.c
```

ภาพที่ 10: แสดงผลลัพธ์การตรวจการคัดลอกด้วย CPD

3.5 ระบบการสอบ

เพื่อป้องกันการทุจริตในการนำระบบตรวจแบบฝึกหัดการเขียนโปรแกรมแบบอัตโนมัติมาใช้ในการสอบ ระบบการสอบ

จึงประกอบด้วยข้อกำหนดดังนี้ (ก) ผู้เข้าสอบใช้งานคอมพิวเตอร์ที่กำหนดไว้เท่านั้น โดยระบบจะตรวจสอบจากหมายเลขไอพีของเครื่องที่เข้าใช้งาน หากไม่ใช่หมายเลขไอพีที่กำหนดจะไม่อนุญาตให้ลงชื่อเข้าใช้ เพื่อเป็นการป้องกันการเข้าถึงระบบจากนอกห้องสอบในระหว่างการสอบ และ (ข) ผู้เข้าสอบจะต้องใช้คอมพิวเตอร์เครื่องดังกล่าวจนเสร็จสิ้นการสอบ โดยบันทึกหมายเลขไอพีที่ใส่ลงชื่อเข้าใช้งานครั้งแรก หากหมายเลขไอพีของผู้เข้าสอบมีการเปลี่ยนแปลงระหว่างการสอบจะถือว่ามีการทุจริตเกิดขึ้น เพื่อเป็นการป้องกันไม่ให้ผู้สอบมอมรับสิทธิ์ผู้ใช้ของตนแก่ผู้อื่นเพื่อให้ส่งโปรแกรมขึ้นรับการตรวจแทนผู้สอบได้ หากคอมพิวเตอร์ที่ใช้ในการสอบเกิดการชำรุดระหว่างการสอบ ผู้สอบสามารถแจ้งผู้คุมสอบหรือผู้ดูแลระบบเพื่อให้ย้ายที่นั่งสอบและล้างค่าการจดจำหมายเลขไอพีเพื่อให้สามารถลงชื่อเข้าใช้งานใหม่ได้

4. ผลการดำเนินงาน

เมื่อเปรียบเทียบกับระบบ Café-Grader แล้วระบบนี้มีความแตกต่างกับระบบเดิมดังตารางที่ 1

ตารางที่ 1: ตารางการเปรียบเทียบระบบเดิมและระบบใหม่

รูปแบบการทำงาน	ระบบเดิม	ระบบใหม่
ภาษาที่รองรับ	ซี	ซี ไพทอน และ แรปเตอร์
ระบบตรวจการคัดลอก	ไม่มี	มีสองระบบ
ระบบบริหารจัดการ โจทย์ และชุดข้อมูลทดสอบ	ทำผ่านไฟล์	ทำผ่านเว็บไซต์
รายงานคะแนน	ทำผ่าน คอมมานไลน์	ทำผ่านเว็บไซต์
ระบบการสอบ	ไม่มี	มี
ระบบตรวจคำบังคับ / คำ ต้องห้ามในรหัสต้นฉบับ โปรแกรม	ไม่มี	มี

จากตารางที่ 1 ระบบที่นำเสนอจะสามารถใช้งานได้สะดวกกว่าเนื่องจากสามารถบริหารจัดการ โจทย์และชุดข้อมูลทดสอบได้ผ่านเว็บไซต์โดยระบบเดิมจะต้องบริหารจัดการ โดยการสร้างไฟล์ขึ้นมาในระบบ รวมถึงระบบเดิมไม่มีระบบตรวจความคล้ายคลึงของโปรแกรมทำให้ไม่สามารถทราบได้ว่าผู้ใดคัดลอกโปรแกรมของผู้อื่นขึ้นมาส่งในระบบ โดยระบบใหม่นั้นมีการตรวจสอบทั้งหมดสองขั้นตอน ทำให้สามารถ

ตรวจสอบความน่าจะเป็นในการคัดลอกโปรแกรมคำตอบของผู้เรียนได้อีกทั้งระบบใหม่ยังรองรับภาษาจำนวนมากกว่าระบบเดิมโดยแบ่งเป็นภาษาขั้นต่ำ (ภาษาซี) ภาษาขั้นสูง (ภาษาไพทอน) และโฟลว์ชาร์ต (ภาษาเรปเตอร์)

ระบบนี้ได้ถูกทดสอบโดยนักศึกษาและอาจารย์ จากนั้นให้ทำการประเมินผ่านแบบสอบถาม โดยมีกลุ่มตัวอย่างเป็นนักศึกษาจำนวน 52 คน อาจารย์จำนวน 4 คน สรุปการประเมินได้ดังนี้

4.1 การประเมินความพึงพอใจของผู้เรียน

ด้านการประเมินประกอบด้วย 4 ด้านดังตารางที่ 2 ซึ่งให้ระดับคะแนนดังต่อไปนี้ มากที่สุด (5) มาก (4) ปานกลาง (3) น้อย (2) และน้อยที่สุด (1)

ตารางที่ 2: สรุปผลการประเมินจากกลุ่มตัวอย่างนักศึกษา

ประเภท	คะแนนเฉลี่ย
ความยากง่ายในการใช้งาน	4.12
ความสวยงามของส่วนติดต่อผู้ใช้	3.87
ความแม่นยำในการตรวจโปรแกรม	3.62
ความพึงพอใจโดยรวม	3.85

4.2 การประเมินความพึงพอใจของผู้สอน

ด้านการประเมินประกอบด้วย 6 ด้านดังตารางที่ 3 ซึ่งให้ระดับคะแนนดังต่อไปนี้ มากที่สุด (5) มาก (4) ปานกลาง (3) น้อย (2) และน้อยที่สุด (1)

ตารางที่ 3: สรุปผลการประเมินจากกลุ่มตัวอย่างนักศึกษา

ประเภท	คะแนนเฉลี่ย
ความยากง่ายในการจัดการ โจทย์	4.25
ระบบตรวจการคัดลอก	4.00
ระบบการสอบ	4.25
ความแม่นยำในการตรวจโปรแกรม	4.00
ระบบการเรียกดูรายงาน	3.75
ความพึงพอใจโดยรวม	4.25

5. สรุป

การใช้ระบบตรวจแบบฝึกหัดการเขียนโปรแกรมแบบอัตโนมัติทำให้ผู้เรียนสามารถตรวจคำตอบของแบบฝึกหัดที่ตนเองทำได้เอง ส่งผลให้เกิดความต่อเนื่องทางความคิดในการพัฒนาโปรแกรมและผู้สอนสามารถบริหารจัดการ โจทย์ ชุด ข้อมูลนำเข้าทดสอบและคำตอบได้อย่างสะดวกรวดเร็ว ส่งผลให้ผู้สอนสามารถพัฒนาคุณภาพของ โจทย์แบบฝึกหัดให้ผู้เรียน

เกิดการเรียนรู้ได้มากขึ้น โดยข้อจำกัดของระบบนี้คือผู้เรียนและผู้สอนจะต้องมีคอมพิวเตอร์ที่สามารถเชื่อมต่อและใช้งานเครือข่ายอินเทอร์เน็ตได้และยังไม่สามารถรองรับการใช้งานพร้อมกันจากผู้ใช้งานจำนวนมากเนื่องจากการตรวจจะต้องตรวจทีละโปรแกรม โดยสรุปผลการประเมินความพึงพอใจต่อระบบโดยรวมของนักศึกษาคือ 3.85 ซึ่งอยู่ในเกณฑ์ค่อนข้างมากและของอาจารย์คือ 4.25 ซึ่งอยู่ในเกณฑ์มาก

6. แนวทางการพัฒนาต่อยอด

6.1 เว็บแอปพลิเคชัน

เพื่อให้สามารถรองรับการใช้งานในสถานศึกษาซึ่งอาจมีการเปิดสอนหลายวิชาพร้อมกัน จึงพัฒนาระบบหลักสูตรโดยแต่ละหลักสูตรจะมีโจทย์เป็นของตนเอง

6.2 ระบบการตรวจ

เพื่อให้สามารถรองรับการตรวจโปรแกรมจำนวนมาก จึงมีแนวคิดพัฒนาการตรวจแบบฝึกหัดแบบกระจายโดยแบ่งภาระการตรวจไปยังเครื่องแม่ข่ายหลายเครื่อง ส่งผลให้เกิดความรวดเร็วในการตรวจและสามารถรองรับการร้องขอการตรวจพร้อมกันจากผู้ใช้งานจำนวนมากได้

6.3 ระบบแซนด์บ็อกซ์

เพื่อให้มีการจำกัดความปลอดภัยในไปทิศทางเดียวกันทุกภาษา จึงจะพัฒนาให้สามารถตรวจภาษาการเขียนโปรแกรมได้หลายภาษาในแซนด์บ็อกซ์เดียว

เอกสารอ้างอิง

- [1] R. Rivest, "The MD5 Message-Digest Algorithm" (Online), Source: <http://tools.ietf.org/html/rfc1321>, 10 Feb 2014.
- [2] D. Eastlake, and P. Jones, "US Secure Hash Algorithm 1 (SHA1)" (Online), Source: <http://tools.ietf.org/html/rfc3174>, 10 Feb 2014.
- [3] J. Fakcharoenphol, "Café-Grader" (Online), Source: <https://github.com/jittat/cafegrader-web>, 4 Dec 2013.
- [4] InfoEther, "CPD Reports" (Online), Source: <http://pmd.sourceforge.net/pmd-5.1.0/cpd.html>, 31 Jan 2014.
- [5] InfoEther, "PMD" (Online), Source: <http://pmd.sourceforge.net>, 31 Jan 2014.
- [6] M. C. Carlisle, "RAPTOR – Flowchart Interpreter" (Online), <http://raptor.martincarlisle.com>, 25 Jan 2014.

โมเดลบอกสเกลเวลาและน้ำหนักเพื่อฝึกทักษะสำหรับโรงเรียนสอนคนตาบอดภาคเหนือใน พระบรมราชินูปถัมภ์ จังหวัดเชียงใหม่

Voice Enabled Weight and Time Apparatuses on Practice for the Northern School for the Blind under the Patronage of the Queen

ยุพดี หัตถลลิน (Upady Hatthasin)¹, สิริพงษ์ มาตาม (Siripong Matame)², และสราวิทย์ ทิมอิม (Sarawit Timim)³
สาขาวิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา ภาคพายัพ เชียงใหม่
UHT@rmutl.ac.th¹

บทคัดย่อ

งานวิจัยนี้ออกแบบประยุกต์ขึ้นเพื่อเป็นอุปกรณ์ช่วยสอนด้านการใช้สเกล ในวิชาที่สอนเกี่ยวกับการบอกเวลาและน้ำหนักบนตาชั่ง สำหรับนักเรียนผู้พิการทางสายตาให้สามารถคลำเพื่อรู้-เข้าใจวิธีการอ่านค่าและมีทักษะการใช้สเกลในโลกของชีวิตจริง เป็นอุปกรณ์ LAB ให้ฝึกหัดพร้อมตรวจสอบคำตอบด้วยการกดปุ่มฟังคำตอบ ช่วยลดภาระผู้สอนที่ต้องพูดประโยคเดิมทุกครั้งและช่วยส่งเสริมให้นักเรียนสามารถทบทวนทักษะได้โดยการพึ่งพาตนเอง หลักการทำงานของอุปกรณ์คือการใช้ไมโครคอนโทรลเลอร์ ตรวจเช็คจุดสัมผัสของแต่ละตัวเลขและควบคุมการทำงาน กรณีนานาฬิกา เมื่อนักเรียนคลำที่เข็มแล้วหมุนเพื่อปรับเวลา จากนั้นใช้อีกมือกดที่ปุ่มบอกเสียง ก็จะมีเสียงบรรยายถึงรายละเอียดของเวลาหน่วยเป็นชั่วโมงและหน่วยนาที ส่วนกรณีตาชั่ง เมื่อนักเรียนวางวัตถุบนตาชั่ง แล้วเอามือคลำที่เข็มสเกลน้ำหนักของตาชั่ง โดยใช้มืออีกมือกดที่ปุ่มบอกเสียง ก็จะมีเสียงบรรยายถึงรายละเอียดของน้ำหนักหน่วยเป็นกิโลกรัมและหน่วยกรัม เสียงที่เปล่งมีทั้งภาษาไทยและภาษาอังกฤษ มีการทดสอบโมเดลสเกลจากผู้มีสายตาดีจำนวน 30 คน และนักเรียนผู้พิการทางสายตากับครูผู้สอนจำนวน 30 คนจากชั้น ป.1 ถึงชั้น ป.3 ผลทดสอบความแม่นยำจากคนสายตาดีอยู่ที่ 86% ของโมเดลบอกเวลาและ 67% ของโมเดลบอกน้ำหนัก ส่วนจากนักเรียนผู้พิการกับครูผู้สอนจะเป็นผลทดสอบความสะดวกในการใช้งาน ซึ่งผลลัพธ์อยู่ที่ 76% ของโมเดลบอกเวลาและ 83% ของโมเดลบอกน้ำหนัก

คำสำคัญ: ผู้พิการทางสายตา นาฬิกาแบบเข็ม เครื่องชั่ง
น้ำหนักแบบเข็ม โรงเรียนสอนคนตาบอด การสัมผัสด้วยมือ

Abstract

This research is designed and applied apparatuses to help on blind students in the course about telling time and weight on the scales. These apparatuses let the blind students feel to know - understand how to read the scale and skill in the world of real life as an intern. They are in laboratory or LAB to help practice and check student's answers by pressing a button to hear the answer. They can reduce the burden on teachers to say the sentence as redundant solutions. They help encourage students to review skills through self-reliance. The process principle of these devices is two cases. In case of the clock, when the student palpates on the clockwise and then rotates to adjust the time, he or she uses the other hand to press the button of voice. There will describe the details of the time in hours and minutes. In case of the weight scales, when the student places objects on the weight scales, and then palpates at the tongue of the weight scales, he or she uses the other hand to press the button of voice. There will be a detailed description of the weight in kilograms and grams. For both of processes, the microcontroller can control in emitting sounds in both Thai and English language. A scale model testing is based on the 30 eyesight peoples and the 30 approximately blind students from grade 1 to grade 3. From the eyesight, the test results are accuracy of 86% in the clock models and the 67% in the weight scales. From the teacher of blind students and the blinds, the examination on the ease of use is based. The evaluation results at 76% and 83% of clock models and the weight scales, respectively.

Keywords: Blind man, Clock, Scales, The Blind School, Hand Touch.

1. บทนำ

จากที่ได้ไปศึกษาที่โรงเรียนสอนคนตาบอดภาคเหนือในพระบรมราชินูปถัมภ์[1] จังหวัดเชียงใหม่ ได้สังเกตเห็นถึงปัญหาที่เกิดขึ้นตรงที่ ขบวนการสอนมีลักษณะสอนแบบตัวต่อตัว โดยผู้สอนจะจับมือนักเรียนไปคลำสัมผัสที่อุปกรณ์สื่อการสอน และในแต่ละบทเรียนอาจารย์ผู้สอนต้องพูดซ้ำหลายๆ ครั้งเพื่อให้ให้นักเรียนได้เข้าใจในเนื้อหาวิชา โดยเฉพาะอย่างยิ่งวิชาที่เกี่ยวกับการเรียนรู้เรื่องสเกล จะมีการนำเครื่องชั่งน้ำหนักจริงขนาด 3 กิโลกรัม และน้ำหนักแบบเข็มจริงที่ใช้แขวนอยู่ตามผนังบ้านเรือนทั่วไป มาเป็นอุปกรณ์การสอนเพื่อให้นักเรียนได้รู้จักโลกความจริงเรื่องสเกล

แล้วมีการให้นักเรียนได้เรียนรู้การอ่านสเกลน้ำหนักของตาชั่งและสเกลเวลาบนเข็มหน้าปัดด้วยการใช้มือนักสัมผัสและฟังเสียงได้ทุกครั้งที่ต้องการ โดยไม่ต้องรบกวนอาจารย์ผู้สอนให้พูดซ้ำหลายครั้ง จึงมีการวิจัยโดยใช้ความรู้ด้านการเขียนโปรแกรมไมโครคอนโทรลเลอร์ด้วยภาษาซี มาควบคุมกลไกของเครื่องตาชั่งและของนาฬิกาที่ซึ่งเป็นอุปกรณ์วัสดุสื่อการสอนสำหรับนักเรียนผู้ตาบอด อีกทั้งมีการพัฒนาให้เหมาะสมสอดคล้องกับเนื้อหาเรื่องสเกลที่เป็นสื่อการสอนอันมีพื้นฐานบนเครื่องชั่งน้ำหนักแบบเข็มและนาฬิกาแบบเข็ม

ด้วยเหตุดังกล่าว งานวิจัยนี้จึงเกิดขึ้นเพื่อให้นักเรียนผู้ตาบอดได้ฝึกการใช้มือนักสัมผัสรับรู้ตำแหน่งเข็ม แล้วสามารถฟังเสียงเฉลยตำแหน่งของเข็มจากลำโพง และส่งผลทำให้นักเรียนผู้ตาบอดสามารถเรียนรู้ จดจำ และสามารถทบทวนฝึกทักษะความรู้ได้ด้วยตนเอง

2. วรรณกรรมที่เกี่ยวข้อง

บุคคลตาบอดแต่กำเนิด[2] นั้นสามารถที่จะตรวจจับข้อมูลทางการสัมผัสได้ดีกว่าคนที่มิใช่ตาบอด จากการศึกษาด้านการรับรู้ที่พิเศษเกิดขึ้นจากความผิดปกติของร่างกายในแง่ของการที่คนตาบอดจะรับรู้ไวกว่าคนปกติ การค้นพบได้ทราบข้อมูลว่า สมอจะมีการปรับตัวเข้ากับการขาดหายไปของการรับรู้ภาพด้วยการเพิ่มความสามารถในการรับรู้-การสัมผัส ดังนั้นจึงจำเป็นต้องมีสื่อการเรียนรู้อันใช้สำหรับการรับรู้-การ

สัมผัส ซึ่งศิริพร[1] ได้ส่งเสริมการทำสื่อการสอนเพื่อให้นักเรียนได้ใช้มือสัมผัสแล้วจินตนาการตามที่ครูบรรยาย ซึ่งครูจะต้องมีความรู้เกี่ยวกับสื่อการสอนและต้องอาศัยสื่อประยุกต์หลายชนิด[3-6] เพื่อช่วยแบ่งเบาภาระการพูดทุกครั้งของครู

สื่อแรกเพื่อการเรียนการสอนระบบสุริยะจักรวาลสำหรับผู้ที่มีความบกพร่องทางการมองเห็น[3] สามารถบอกรายละเอียดต่างๆ ของดวงดาวได้ เช่น ขนาดของดวงดาว, ภูมิภาค, ลักษณะเด่น เป็นต้น ทำงานได้สองโหมดคือ: 1) โหมดบรรยาย โดยสื่อจะเปล่งเสียงบรรยายลักษณะของดวงดาวเมื่อมีการสัมผัสแสงของเซนเซอร์ มีการบรรยายได้ทั้งภาษาไทยกับภาษาอังกฤษ 2) โหมดทดสอบ เมื่อมีการสัมผัสแสงของเซนเซอร์ สื่อจะไม่เปล่งเสียงแต่ระบบแสงทำงานตามปกติต้องใช้คู่กับการตั้งคำถามเพื่อให้นักเรียนเลือกดวงดาวด้วยการสัมผัสเพื่อตอบคำถาม

สื่อที่สองเพื่อสื่อการสอนอักษรเบรลล์ขั้นพื้นฐาน สำหรับผู้ที่มีความพร่องทางการเห็น[4] ถูกออกแบบเพื่อผู้ที่มีความพร่องทางการเห็นใช้งานโดยเฉพาะ มีการทำงานสองโหมดคือ โหมดแป้นปุ่มกดอักษรเบรลล์[5] และโหมดเกมส์ โดยสามารถเสียบหูฟังเพื่อศึกษาด้วยตัวเองได้ มีแหล่งจ่ายไฟอยู่สองส่วนคือ 1) แบตเตอรี่ใช้งานได้นานกว่า 10 ชั่วโมง และ 2) ปลั๊กไฟบ้าน

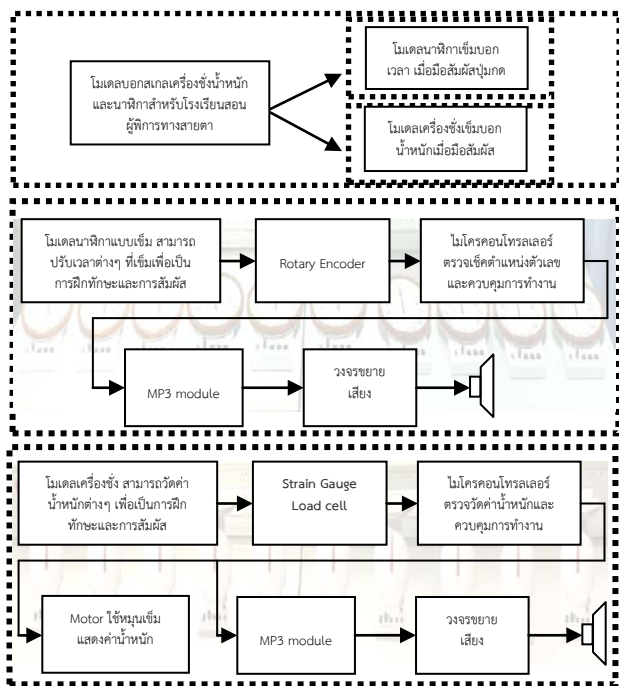
สื่อสุดท้ายเพื่อวัสดุช่วยชี้ตำแหน่งบาร์โค้ดของระบบสื่อการสอนเพื่อให้นักเรียนผู้พิการทางสายตา[6] มีการประยุกต์และพัฒนาระบบช่วยผู้พิการทางสายตาในการเรียนการสอนได้ในหลายๆ วิชาที่เกี่ยวกับการบรรยาย มีการบอกเนื้อหาเฉพาะประจำสื่อชิ้นๆ หรือวิชาที่จะใช้เรียน มีการคิดรหัสบาร์โค้ดบนวัสดุชี้ตำแหน่งเพื่อให้นักเรียนสามารถคลำหาตำแหน่งเป้าหมายแล้วยังสัญญาณด้วยเครื่องอ่านบาร์โค้ดได้ด้วยตนเองช่วยเป็นอุปกรณ์ข้อมูลสื่อให้แก่ครูในการเตรียมสอนตามเป้าหมายว่า ข้อมูลสื่อคืออะไร ชนิดใด เรื่องใด ดังนั้นจึงช่วยอำนวยความสะดวกแก่ทั้งครูและนักเรียนผู้พิการทางสายตาในการศึกษา อีกทั้งสื่อนี้จะช่วยให้ผู้ที่ดูแลผู้พิการสามารถทราบสถิติได้ว่า มีการใช้บริการเกี่ยวกับสื่อการสอนชนิดใดมากที่สุด สำหรับเป็นข้อมูลในการจัดเตรียมหรือผลิตสื่อการสอนชนิดนั้นๆ ให้มีจำนวนที่เหมาะสมในลำดับต่อไป

งานวิจัยดังกล่าวได้เอื้อประโยชน์ด้านการศึกษาให้แก่ นักเรียนผู้มีความพร่องทางการเห็น และช่วยลดภาระครูผู้สอนที่

ต้องพูดสอนซ้ำหลายๆครั้ง ได้เป็นอย่างดี แต่ก็ยังมีอีกหลายสิ่ง
ที่นักเรียนจำเป็นต้องได้เรียนรู้ในโลกชีวิตจริง โดยเฉพาะอย่างยิ่ง
การเรียนรู้เกี่ยวกับสเกลของเวลาและสเกลของน้ำหนักจาก
ตาชั่ง ซึ่งเป็นงานวิจัยนวัตกรรมเชิงประยุกต์ใหม่

3. วิธีการดำเนินการวิจัย

จากที่ได้ศึกษาหลักการทำงานของนาฬิกา[7] และเครื่องชั่ง
น้ำหนัก[8] เพื่อเป็นแนวทางในการประยุกต์สร้างโมเดลนาฬิกา
แบบเข็มบอกเวลาและโมเดลเครื่องชั่งน้ำหนักแบบเข็มบอก
น้ำหนัก แสดงเป็นกรอบแนวคิดประยุกต์ดังในภาพที่ 1



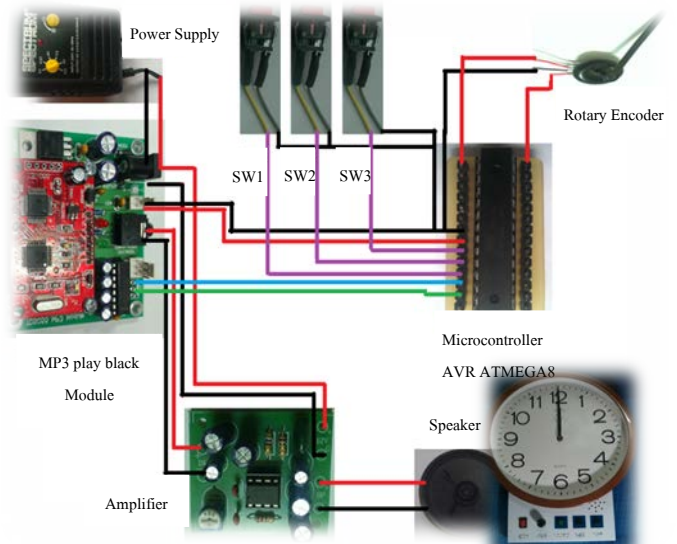
ภาพที่ 1: กรอบแนวคิดการออกแบบประยุกต์

มีการใช้ไมโครคอนโทรลเลอร์[9-10] ตรวจสอบเช็คจุดสัมผัส
ของแต่ละตัวเลขและควบคุมการทำงาน เมื่อปรับค่าที่เข็มแล้ว
สัมผัสที่ปุ่มก็จะมีเสียงบรรยายถึงรายละเอียดของค่าตัวเลขของ
ตาชั่งและเวลาของนาฬิกาจากวงจรบันทึกเสียงบรรยาย
รายละเอียดของแต่ละตัวเลขและเพื่อให้ผู้เรียนได้สามารถใช้
อุปกรณ์นี้ มีการออกแบบสองส่วนตามลำดับหัวข้อต่อไปนี้

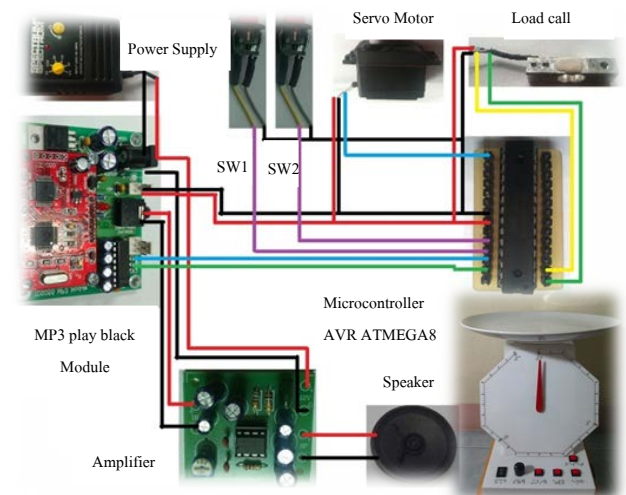
3.1 การออกแบบและสร้างโมเดลนาฬิกาเข็มบอกเวลา

เริ่มจากแหล่งจ่ายไฟ 9V_{DC} จ่ายไฟให้กับระบบ
Microcontroller AVR ATMEGA8 ตรวจสอบเช็คตำแหน่งของเข็ม
นาฬิกาจาก Rotary Encoder ที่ขา PB0 และ ขา PB1 แล้ว
ตรวจสอบเช็คการกดฟังเสียงเวลาภาษาไทยจาก SW1(Thai Button)

ที่ขา PD2 ตรวจสอบเช็คการกดฟังเสียงเวลาภาษาอังกฤษจาก
SW2(Eng Button) ที่ขา PD3 และเวลาเป็น AM หรือ PM จาก
SW3(AM or PM)ที่ขา PD4 หลังจากนั้นส่งให้ MP3 Module
ทำงานผ่าน Serial Port ขา PD0 และ ขา PD1 ทำการขยายเสียง
โดย วงจร Amplifier 2W และส่งเอาท์พุตเสียงออกไปยังลำโพง
ดังแสดงในภาพที่ 2



ภาพที่ 2: การเชื่อมต่ออุปกรณ์ของโมเดลนาฬิกาแบบเข็มมีเสียงบอกเวลา



ภาพที่ 3: การเชื่อมต่ออุปกรณ์ของโมเดลเครื่องชั่งน้ำหนักแบบเข็มมีเสียงบอกน้ำหนัก

3.2 การออกแบบและสร้างโมเดลเครื่องชั่งมีเสียงบอกน้ำหนัก

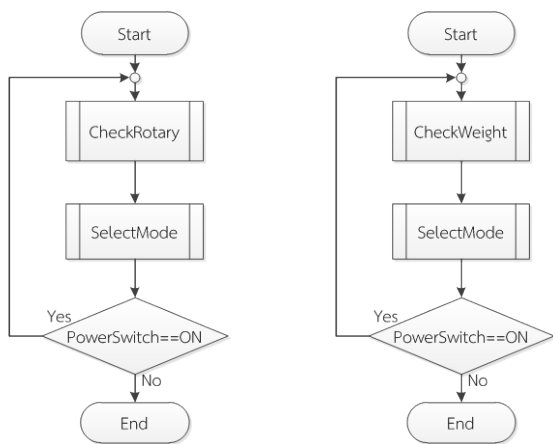
เริ่มจากแหล่งจ่ายไฟ 9V_{DC} จ่ายไฟให้กับระบบ
Microcontroller AVR ATMEGA8 ตรวจสอบเช็คน้ำหนักจาก Strain
Gauge Load cell ที่ขา PC5 และสั่งให้ Servo Motor หมุนไปยัง
ตำแหน่งตามน้ำหนักที่วัดได้ แล้วตรวจสอบเช็คการกดฟังเสียงเวลา

ภาษาไทยจาก SW1(Thai Button) ที่ขา PD2 ตรวจสอบเช็คการกดฟังเสียงเวลาภาษาอังกฤษจาก SW2(Eng Button) ที่ขา PD3 หลังจากนั้นสั่งให้ MP3 Module ทำงานผ่าน Serial Port ขา PD0 และ ขา PD1 ทำการขยายเสียงโดย วงจร Amplifier 2W และส่งเอาต์พุตเสียงออกไปยังลำโพง ดังแสดงในภาพที่ 3

3.3 การออกแบบและสร้างในส่วนของซอฟต์แวร์

ในภาพที่ 4.(ก) เป็นโฟลว์ชาร์ตการทำงานของโปรแกรมโมเดลนาฬิกา การทำงานระบบเริ่มจากการเรียกใช้ฟังก์ชันย่อย CheckRotary แล้วเรียกใช้ฟังก์ชันย่อย SelectMode แล้วเช็ค Power Switch ถ้ายังมีการ ON ให้ทำคำสั่งไปเรื่อยๆ

ฟังก์ชัน CheckRotary เป็นการเช็คตำแหน่งการหมุนของเข็มนาฬิกาโดยใช้ Rotary Encoder การทำงานจะเช็คว่า ถ้ามีการหมุนไปด้านขวา ก็จะทำให้ Clock ทำการบวกค่าเพิ่ม และถ้ามีการหมุนไปด้านซ้าย ก็จะทำให้ Clock ทำการลดค่าลง



ก) โมเดลนาฬิกา ข) โมเดลเครื่องชั่งน้ำหนัก

ภาพที่ 4 : โฟลว์ชาร์ตการทำงานของโปรแกรม

ฟังก์ชัน SelectMode เป็นการตรวจเช็คโหมดเลือกรับฟังเสียง โดยจะทำการตรวจเช็ค สวิตช์ ThaiButton ถ้ามีการกดให้ไปทำงานที่ฟังก์ชันย่อย ThaiVoice เป็นบอกเวลาเป็นภาษาไทย และตรวจเช็คสวิตช์ EngButton เพื่อบอกเวลาเป็นภาษาอังกฤษ ถ้าไม่มีการกดปุ่ม ก็ให้จบการทำงานฟังก์ชันย่อย SelectMode

ในภาพที่ 4.(ข) เป็นโฟลว์ชาร์ตการทำงานของโปรแกรมโมเดลเครื่องชั่งน้ำหนัก การทำงานระบบเริ่มจากการเรียกใช้ฟังก์ชันย่อย CheckWeight และเรียกใช้ฟังก์ชันย่อย

SelectMode และเช็ค Power Switch ถ้ายังมีการ ON ก็ให้ทำคำสั่งไปเรื่อยๆ

ฟังก์ชัน CheckWeight เป็นการเช็คค่าน้ำหนักจากวัตถุที่ชั่งโดยใช้โวลต์เซลล์เซนเซอร์แล้วคำนวณค่าจากสัญญาณอนาล็อกที่เข้ามาและแปลงเป็นดิจิตอลเพื่อให้ได้ค่าน้ำหนัก

ฟังก์ชัน SelectMode เป็นการตรวจเช็คโหมดเลือกรับฟังเสียง โดยจะทำการตรวจเช็ค สวิตช์ ThaiButton ถ้ามีการกดให้ไปทำงานที่ฟังก์ชันย่อย ThaiVoice เป็นบอกน้ำหนักเป็นภาษาไทย และตรวจเช็คสวิตช์ EngButton เพื่อบอกน้ำหนักเป็นภาษาอังกฤษ ถ้าไม่มีการกดปุ่ม ก็ให้จบการทำงานฟังก์ชันย่อย SelectMode

4. ผลการดำเนินงาน

หลังจากที่ได้ออกแบบวงจรและประกอบเข้าเป็นตัวเครื่องตามขั้นตอนการสร้างแล้ว ก็เข้าสู่การทดสอบเกี่ยวกับวิธีการใช้งาน คุณสมบัติและขีดจำกัดของอุปกรณ์ที่นำเสนอ เพื่อเป็นข้อมูลเบื้องต้นสำหรับผู้ที่จะนำไปศึกษาเพื่อสร้างใช้งานและปรับปรุงในเชิงอุตสาหกรรมหรือนำไปพัฒนาให้มีคุณสมบัติให้ดียิ่งขึ้นต่อไป มีการทดลองดังนี้

4.1 วิธีการทดลองกับทั้ง 2 โมเดล

มีการสร้างแบบทดสอบให้คนสายตาปกติ 30 คนได้ทดสอบและประเมินผลในช่วงแรก แล้วจึงให้นักเรียนผู้พิการทางสายตาและครูผู้สอน 30 คนในชั้น ป.1 ถึงชั้น ป.3 ได้ทำการทดสอบใช้งานโมเดลนาฬิกาแบบเข็มมีเสียงบอกเวลาและทำแบบประเมินดังในภาพที่ 5 และทดสอบใช้งานโมเดลเครื่องชั่งน้ำหนักแบบเข็มมีเสียงบอกน้ำหนักและทำแบบประเมินดังในภาพที่ 6 แนวการประเมินมี 5 ประเด็นคือ ด้านความปลอดภัย ด้านการออกแบบ ด้านความแม่นยำ ด้านความสะดวกในการใช้งาน และด้านราคา

4.2 สรุปผลรวมการทดสอบทั้งผู้มีสายตาปกติและครูผู้สอนกับนักเรียนผู้พิการทางสายตา

มีการนำเอาผลประเมินของทั้ง 2 โมเดลมาวิเคราะห์แล้วเปรียบเทียบซึ่งกันและกันระหว่างผู้ที่มีสายตาปกติและผู้พิการทางสายตา โดยจะเป็นการบรรยายในลำดับต่อไปนี้

กรณีโมเดลนาฬิกาแบบเข็มมีเสียงบอกเวลา: เมื่อนำเอาผลการประเมินจากผู้มีสายตาปกติและผู้พิการทางสายตามาพล็อตกราฟรวมกันแสดงดังภาพที่ 7 โดยใช้สเกลรวมในหน่วยของ

เปอร์เซ็นต์ ผลจะถูกตีความหมายแสดงให้รู้ว่ามี ความแตกต่างกัน ในมุมมองของผู้มีสายตาสกปรกและผู้พิการทางสายตาดังนี้คือ

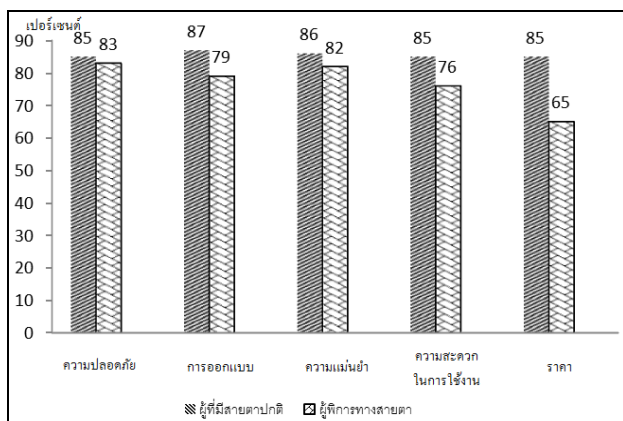
1) ทัศนคติความปลอดภัย ผู้มีสายตาสกปรกและผู้พิการทาง



ภาพที่ 5: นักเรียนผู้พิการทางสายตา กำลังทดสอบใช้งานโมเดลนาฬิกาแบบเข็มมีเสียงบอกเวลา



ภาพที่ 6: นักเรียนผู้พิการทางสายตา กำลังทดสอบใช้งานโมเดลเครื่องชั่งน้ำหนักแบบเข็มมีเสียงบอกน้ำหนัก



ภาพที่ 7: กราฟผลการประเมินของโมเดลนาฬิกาบอกเวลาจากผู้ทดสอบที่สายตาสกปรกและผู้พิการทางสายตา

สายตาให้ผลประเมินใกล้เคียงกัน หมายความว่าไม่มีทัศนคติแบบเดียวกัน

2) ทัศนคติการออกแบบ ผู้มีสายตาสกปรกให้คะแนนมากกว่า 6 คะแนน ซึ่งหมายความว่าผู้มีสายตาสกปรกใช้สายตาเป็นตัวประเมินซึ่งมีความน่าเชื่อถือได้

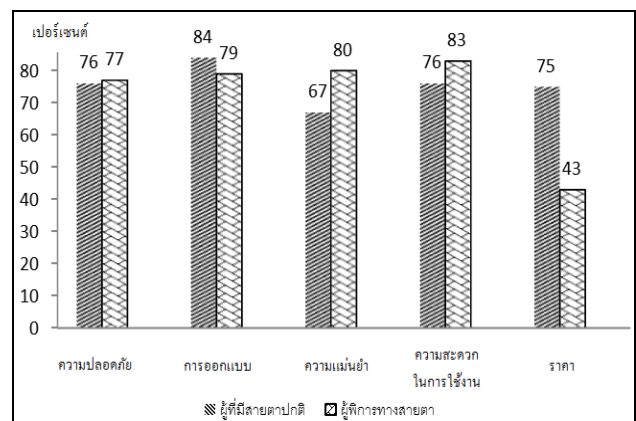
3) ทัศนคติความแม่นยำ ผู้มีสายตาสกปรกให้คะแนนมากกว่า 4 คะแนนซึ่งหมายความว่าผู้มีสายตาสกปรกได้ใช้เครื่องมือที่สามารถมองเห็นได้ด้วยสายตาตัวเอง ด้วยเหตุนี้ผลประเมินจึงอยู่ในระดับน่าพอใจ

4) ทัศนคติความสะดวกในการใช้งาน ผู้มีสายตาสกปรกให้คะแนนมากกว่า 9 คะแนนเนื่องเพราะถึงแม้เครื่องมือมีภาษาเบรลล์ก็ยังมีสัญลักษณ์ซึ่งทำให้ผู้มีสายตาสกปรกสามารถใช้เครื่องมือได้สะดวกขึ้น

5) ทัศนคติราคา ผู้พิการทางสายตาให้คะแนนน้อยกว่าถึง 20 คะแนน เพราะผู้พิการทางสายตายังขาดประสบการณ์ในการประเมินราคาเนื่องจากมีอายุอยู่ในช่วงชั้น ป.1 ถึงชั้น ป.3 จึงทำให้คะแนนแตกต่างกันพอสมควร ด้วยเหตุนี้เกณฑ์ราคาจึงสมเหตุสมผลจริง

กรณีโมเดลเครื่องชั่งน้ำหนักแบบเข็มมีเสียงบอกน้ำหนัก: เมื่อนำเอาผลการประเมินจากผู้มีสายตาสกปรกและผู้พิการทางสายตามาพล็อตกราฟรวมกันดังภาพที่ 8 โดยใช้สเกลรวมในหน่วยเปอร์เซ็นต์ ผลจะถูกตีความหมายให้รู้ว่ามี ความแตกต่างกัน ในมุมมองของผู้มีสายตาสกปรกและผู้พิการทางสายตาดังนี้คือ

1) ทัศนคติความปลอดภัย ผู้ที่มีสายตาสกปรกและผู้พิการทางสายตาให้ผลประเมินใกล้เคียงกัน ซึ่งหมายความว่าไม่มีทัศนคติแบบเดียวกัน



ภาพที่ 8: กราฟผลการประเมินของโมเดลเครื่องชั่งน้ำหนักจากผู้ทดสอบที่สายตาสกปรกและผู้พิการทางสายตา

2) กรณีการออกแบบ ผู้มีสายตาปกติให้คะแนนมากกว่า ซึ่งหมายความว่าผู้มีสายตาปกติใช้สายตาเป็นตัวประเมินผล ซึ่งมีความน่าเชื่อถือได้

3) กรณีความแม่นยำ ผู้พิการทางสายตาให้คะแนนมากกว่า 13 คะแนนซึ่งหมายความว่าผู้พิการทางสายตาได้ใช้เครื่องมือเปิดโลกทัศน์ให้กับตัวเอง ดังนั้นผลประเมินจึงอยู่ในระดับน่าพอใจ

4) กรณีความสะดวกในการใช้งาน ผู้พิการทางสายตาให้คะแนนมากกว่า 8 คะแนนเนื่องจากเครื่องมือมีภาษาเบรลล์ซึ่งทำให้ผู้พิการทางสายตาสามารถใช้เครื่องมือได้สะดวกขึ้น

5) กรณีราคา ผู้พิการทางสายตาให้คะแนนน้อยกว่าถึง 32 คะแนนเพราะยังขาดประสบการณ์ในการประเมินราคาด้วยมีอายุอยู่ในช่วงชั้น ป.1 ถึง ป.3 จึงทำให้คะแนนแตกต่างกันพอสมควร ดังนั้นผลประเมินราคาจากผู้มีสายตาปกติ จึงเป็นที่น่าเชื่อถือได้มากกว่า ด้วยเหตุนี้เกณฑ์ราคาจึงสมเหตุสมผลจริง

5. สรุป

งานวิจัยนี้คือการออกแบบประยุกต์และสร้างขึ้นเพื่อเป็นอุปกรณ์ช่วยการเรียนการสอนในชั้นเรียนแก่นักเรียนผู้พิการทางสายตา ในระดับชั้นประถม เพื่อให้พวกเขาสามารถเรียนรู้เรื่องสเกลเวลาและสเกลน้ำหนักในโลกของชีวิตจริง อีกทั้งสามารถฝึกฝน-ทบทวนการเรียนรู้ได้ด้วยตนเอง โดยได้คำนึงถึงผลประโยชน์ของนักเรียนเป็นหลักในแง่ความปลอดภัย ความแม่นยำ ใช้งานง่าย และราคาถูก โมเดลสเกลทั้งสองจะใช้ไมโครคอนโทรลเลอร์ควบคุมให้สามารถแปลงเสียงได้ทั้งภาษาไทยและภาษาอังกฤษ มีผลทดสอบจากคนสายตาดีจำนวน 30 คน และนักเรียนผู้พิการกับครูผู้สอนจำนวน 30 คน ในชั้น ป.1 ถึง ป.3 ผลทดสอบความแม่นยำจากคนสายตาดีอยู่ที่ 86% ของโมเดลบอกเวลาและ 67% ของโมเดลบอกน้ำหนัก เนื่องจาก Strain Gauge Load cell มีราคาปานกลางจึงทำให้มีแม่นยำปานกลาง ส่วนจากนักเรียนผู้พิการกับครูผู้สอนจะเป็นผลทดสอบความสะดวกในการใช้งาน ซึ่งผลลัพธ์อยู่ที่ 76% ของโมเดลบอกเวลาและ 83% ของโมเดลบอกน้ำหนัก เนื่องจากโมเดลสเกลทั้งสองนี้มีราคาต้นทุนไม่สูงมากและสามารถตอบสนองด้านการเรียนการสอนได้ดี ดังนั้นจึงมีการจัดสร้างขึ้นเพื่อมอบให้โรงเรียนสอนคนตาบอดภาคเหนือในพระบรมราชินูปถัมภ์ จังหวัดเชียงใหม่ ได้ใช้ประโยชน์ด้าน

การเรียนการสอนเป็นจำนวน 10 เครื่องบอกสเกลเวลา และจำนวน 10 เครื่องบอกค่าสเกลน้ำหนัก

6. กิตติกรรมประกาศ

งานวิจัยนี้ขอขอบคุณท่านอาจารย์ศิริพร วงศ์จุฬโรจน์[1] ที่ให้รายละเอียดเกี่ยวกับผู้พิการทางสายตา, ขอขอบคุณสถาบันวิจัยและพัฒนา แห่งมหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา ภาคพายัพ เชียงใหม่และ สำนักงานคณะกรรมการการอุดมศึกษา ที่ได้มอบทุนการวิจัย

เอกสารอ้างอิง

- [1] ศิริพร วงศ์จุฬโรจน์, โรงเรียนสอนคนตาบอดภาคเหนือในพระบรมราชินูปถัมภ์ อำเภอเมือง จังหวัดเชียงใหม่
- [2] นุศลซึ่งตาบอดแต่กำเนิด, [Online Access: 31 ม.ค. 2557], <http://www.vcharkarn.com/vnews/153361>
- [3] กันต์ธร ณ ลำพูน และจิตติพัฒน์ บุญ, “สื่อการเรียนการสอนระบบสุริยะจักรวาลสำหรับผู้ที่มีความบกพร่องทางการมองเห็น”, ปรินญาณินพนธ์, สาขาวิชาครุศาสตร์ศึกษาศาสตร์, มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา ภาคพายัพ เชียงใหม่ (มทร.ล้านนา), อ.เมือง จ.เชียงใหม่, 2555.
- [4] จาตุรงค์ แอบคำ และณัฐภูมิ วรพันธ์, “สื่อการสอนอักษรเบรลล์ขั้นพื้นฐานสำหรับผู้ที่มีความบกพร่องทางการเห็น”, ปรินญาณินพนธ์, สาขาวิชาครุศาสตร์ศึกษาศาสตร์, มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา ภาคพายัพ เชียงใหม่, อ.เมือง จ.เชียงใหม่, 2555.
- [5] งานพัฒนาสื่อเทคโนโลยีและนวัตกรรมโรงเรียนสอนคนตาบอดภาคเหนือฯ, “ทักษะการ อ่านเขียนและพิมพ์อักษรเบรลล์ขั้นพื้นฐาน”, โรงเรียนสอนคนตาบอดภาคเหนือในพระบรมราชินูปถัมภ์ อ.เมือง จ.เชียงใหม่, 2547.
- [6] ยุพดี หัตถสิน, ธวัชชัย ตาใจ, และอรรรพณ ณ ตะกั่วทุ่ง, “วัสดุช่วยชี้ตำแหน่งบาร์โค้ดของระบบสื่อการสอนเพื่อนักเรียนผู้พิการทางสายตา”, NCCIT2013, 9-10 พ.ค. 2013, หน้า 940-945.
- [7] ประวัติความเป็นมาของนาฬิกา, [Online Access: 31 ม.ค. 2557], <http://www.siamsouth.com/smf/index.php?topic=6781.0>
- [8] ความรู้ทั่วไปเกี่ยวกับเครื่องชั่ง, [Online Access: 31 ม.ค. 2557], <http://เครื่องชั่งไทย.com/how2use/>
- [9] เอกชัย มะการ, “เรียนรู้และเข้าใจ ไมโครคอนโทรลเลอร์ MCS-51 ด้วยภาษาเบสิก BASCOM-8051”, บ.ทีที จำกัด, 2547, 240 หน้า.
- [10] โอภาส ศิริกรรชิตถาวร, “เรียนรู้และพัฒนาไมโครคอนโทรลเลอร์ ARM7 ด้วยภาษาซี”, บ.อีทีที จำกัด, 2542, 163 หน้า.

ระบบติดตามสถานการณ์สถานศึกษาที่ประสบอุทกภัย โดยใช้สมาร์ตโฟนแอนดรอยด์ A Monitoring System for School Flood Victims Using the Android Smart Phone

มยุรี จีระมาตย์ (Mayuree Jeeramard)¹ และศักดิ์ชาย ตั้งวรรณวิทย์ (Sakchai Tangwannawit)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹my_aoun@hotmail.com, ²sakchait@kmutnb.ac.th

บทคัดย่อ

วัตถุประสงค์ของงานวิจัยนี้ เพื่อพัฒนาระบบติดตามสถานการณ์สถานศึกษาที่ประสบอุทกภัยโดยใช้สมาร์ตโฟนแอนดรอยด์ ซึ่งระบบนี้สามารถรองรับการรายงานข้อมูลอุทกภัยแบบเรียลไทม์ โดยสถาปัตยกรรมของการพัฒนาระบบมี 2 ส่วน ส่วนที่ 1 สมาร์ตโฟนแอนดรอยด์ ใช้ในการรายงานพิกัดจีพีเอส (Global Positioning System: GPS) ของสถานศึกษา ซึ่งสามารถแสดงผลผ่าน Google Map API ส่วนที่ 2 คือ Web-Based Application สามารถรองรับการนำเข้าข้อมูลประกอบด้วย ชื่อสถานศึกษา พิกัดจีพีเอส และรายงานการทำงานของแอปพลิเคชันที่ติดตั้งบนสมาร์ตโฟน จะแสดงตำแหน่งปัจจุบันของผู้ใช้งานบนกูเกิลแมพ โดยผู้ใช้งานสามารถยืนยันตำแหน่งที่ถูกต้อง หรือค้นหาชื่อสถานศึกษาที่ต้องการและรายงาน โดยเพิ่มรูปถ่ายที่เปิดใช้งานจีพีเอสบนสมาร์ตโฟน การเพิ่มข้อมูลสามารถเพิ่มข้อมูลเกี่ยวกับความเสียหาย ระดับน้ำ และรายละเอียดข้อมูลอื่นๆ ได้ ซึ่งสถานศึกษาอื่นนอกจากนี้ยังสามารถยืนยันรายละเอียดและเพิ่มรูปถ่าย ที่รอการยืนยันความถูกต้องกับรหัสสถานศึกษาเพื่อเพิ่มเข้าสู่ฐานข้อมูลผลการทดสอบประสิทธิภาพของระบบ จำนวน 30 แห่ง โดยทดสอบแสดงข้อมูล ค้นหาตำแหน่งปัจจุบันของสถานศึกษาด้วยชื่อ ที่อยู่ และข้อมูลของสถานศึกษาประกอบด้วย ระดับความเสียหาย ระดับน้ำ รูปถ่าย โดยสรุปผลการทดสอบความถูกต้องของระบบที่พัฒนาขึ้น เพื่อติดตามสถานการณ์สถานศึกษาที่ประสบอุทกภัย พบว่ามีความสำเร็จและเป็นไปตามวัตถุประสงค์

คำสำคัญ: จีพีเอส กูเกิลแมพ สมาร์ตโฟน แอนดรอยด์

Abstract

The objective of this research is to develop a monitoring system for school flood victims using the Android smart phone. This system provides the ability to report flood information in real time. The architecture of the developed system has two parts, 1. the Android smart phone is used to report the coordinates via GPS (Global Positioning System) of school which can be displayed over Google maps API. 2, a web-based application which provides the ability to enter data such as school name, GPS coordinates, and the report. The application installed on the smart phone will display the user's current position over Google maps. The user can then verify the correct location or search for the school's name, and report by adding photos with GPS coordinates. Additional information can also be added referring to amount of damage, the water level and other report details. Other schools can also confirm details and add photographs according with school's code to increase the database. The accuracy of the system was tested in 30 locations to test the information displayed such as current position with name, address and information of school; damage level, water level, photos. In conclusion, from the positive results of the accuracy tests, the developed monitoring system for school flood victims clearly achieved all objectives successfully.

Keywords: GPS, Google Map, Smart Phone, Android

1. บทนำ

กระทรวงศึกษาธิการ มีสถานศึกษาและสำนักงานตั้งอยู่ในเขตพื้นที่ทั่วประเทศ จากเหตุการณ์ที่เกิดอุทกภัยในแต่ละครั้ง มีสถานศึกษาและสำนักงานได้รับความเสียหายจำนวนมาก สถานศึกษาจะทำการรายงานข้อมูลความเสียหายที่เกิดจากอุทกภัยให้ส่วนกลางรับทราบ ซึ่งในปัจจุบันนี้ สถานศึกษาเป็นผู้จัดส่งข้อมูลตามเอกสารแบบฟอร์มความเสียหาย โดยส่วนกลางทำการรวบรวมข้อมูลซึ่งมีหลายรูปแบบ เช่น การรับข้อมูลทางโทรสารและรูปแบบไฟล์ เพื่อทำการสรุปข้อมูลในการนำเสนอผู้บริหาร ซึ่งการรายงานข้อมูลในรูปแบบดังกล่าวพบว่า มีความล่าช้าและสถานศึกษามีชื่อเหมือนกันหลายแห่ง ข้อมูลมีการปรับเปลี่ยนตามการรายงานอยู่ตลอดเวลา ทำให้ต้องใช้เวลาในการสรุปข้อมูล และทำให้มีความคลาดเคลื่อนจากการรายงานในเวลาที่ต่างกัน จึงทำให้ผู้บริหารได้รับทราบข้อมูลที่ล่าช้าและคลาดเคลื่อน

โทรศัพท์มือถือ สมาร์ทโฟน สามารถรองรับระบบปฏิบัติการต่างๆ ได้ และระบบปฏิบัติการแอนดรอยด์ เป็นระบบปฏิบัติการแบบเปิดเผยแพร่แวร์ต้นฉบับ (Open Source) ที่สามารถพัฒนาแอปพลิเคชันต่างๆ ได้อย่างกว้างขวาง การเชื่อมต่ออินเทอร์เน็ตผ่านสมาร์ตโฟนยังสามารถระบุพิกัดจีพีเอสจากรูปถ่ายได้ จากคุณสมบัติของสมาร์ตโฟนระบบปฏิบัติการแอนดรอยด์ดังกล่าว และปัญหาในการรายงานข้อมูลสถานศึกษาที่ประสบอุทกภัย จึงทำให้มีแนวคิดในการพัฒนาระบบติดตามสถานการณ์สถานศึกษาที่ประสบอุทกภัยโดยใช้สมาร์ตโฟนระบบปฏิบัติการแอนดรอยด์ เพื่อใช้ในการรายงานสถานการณ์ของสถานศึกษาในสังกัดกระทรวงศึกษาธิการที่ประสบอุทกภัยให้สามารถแสดงตำแหน่งปัจจุบันของสถานศึกษาโดยใช้พิกัดจีพีเอส (Global Positioning System :GPS) และเพิ่มข้อมูลรูปถ่ายที่มีการฝังพิกัดจีพีเอสจากสถานที่จริง พร้อมทั้งระบุระดับความเสียหายของอุทกภัย ระดับน้ำ และยังสามารถใช้พิกัดจีพีเอสจากรูปถ่ายบนสมาร์ตโฟน เพื่อเพิ่มตำแหน่งของสถานศึกษาได้ ทำให้การรายงานสถานการณ์อุทกภัยมีความสะดวก รวดเร็ว และช่วยให้ผู้บริหารใช้ในการประเมินสถานการณ์และให้ความช่วยเหลือสถานศึกษาในเบื้องต้นได้

ทันท่วงที และเป็นฐานข้อมูลเพื่อเตรียมความพร้อมในการรับมือการเกิดอุทกภัยในอนาคตได้

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 จีพีเอส (Global Positioning System : GPS)

บนสมาร์ตโฟนระบบปฏิบัติการแอนดรอยด์

จีพีเอสบนสมาร์ตโฟนระบบปฏิบัติการแอนดรอยด์ โดยทั่วไป มี 2 แบบ คือ GPS และ A-GPS ซึ่ง GPS นั้น จะรับสัญญาณตำแหน่ง (ข้อมูลตำแหน่งและข้อมูลเวลาปัจจุบัน) จากดาวเทียม GPS โดยตรง ส่วน A-GPS (Assisted GPS) จะรับสัญญาณตำแหน่งจากสถานีฐานของเครือข่ายโทรศัพท์มือถือ (GPS base station) ทำหน้าที่รับข้อมูลต่างๆ จากดาวเทียม GPS โดยตรง แล้วค่อยส่งต่อให้กับเครื่องรับที่เป็น A-GPS อีกที จึงทำให้มีข้อดีในเรื่องของความเร็วในการทำงาน แต่จะมีข้อเสียคือไม่สามารถใช้งานได้ในพื้นที่ที่ไม่สามารถรับสัญญาณโทรศัพท์ได้ และ GPS Application Software บางตัวจะไม่สามารถทำงานได้ถ้าไม่มีการเปิดใช้งานอินเทอร์เน็ตร่วมด้วย ซึ่ง A-GPS ส่วนใหญ่จะเป็นโทรศัพท์มือถือ แต่มีบางรุ่นที่เป็น GPS ซึ่งสามารถรับสัญญาณจากดาวเทียมโดยตรงได้ เช่น HTC Desire S, HTC Sensation, HTC Incredible เป็นต้น [1]

2.2 ทฤษฎีเกี่ยวกับกูเกิลแมพ เอพีไอ (Google Map API)

API มาจากคำว่า Application Programming Interface ซึ่งหมายถึงการเขียนโปรแกรมโดยมีการเรียกใช้ Library เช่น งาน Routines, Data Structures, Object class และตัวแปร ดังนั้น Google Map API ก็หมายถึงการเรียกใช้ข้อมูล Library จาก Google ซึ่งเหมาะสำหรับนักพัฒนาโปรแกรมจากภาษา JavaScript บริการ API ของ Google เช่น Google Maps SDK for iOS เขียนโปรแกรมบนที่บนระบบปฏิบัติการ iOS Apple Google Maps Android API v2 เขียนโปรแกรมบนที่บนระบบปฏิบัติการ Android [2]

2.3 ระบบปฏิบัติการแอนดรอยด์ (Android)

ระบบปฏิบัติการแบบเปิดเผยแพร่แวร์ต้นฉบับ (Open Source) โดยบริษัท กูเกิล (Google Inc.) ที่ได้รับความนิยมเป็นอย่างสูง เนื่องจากอุปกรณ์ที่ใช้ระบบปฏิบัติการแอนดรอยด์มีจำนวนมาก อุปกรณ์มีหลากหลายระดับ หลายราคา รวมทั้งสามารถทำงานบนอุปกรณ์ที่มีขนาดเล็กและมีความละเอียด

แตกต่างกันได้ การพัฒนาโปรแกรมเพื่อใช้งานบนระบบปฏิบัติการแอนดรอยด์ มีข้อมูลในการพัฒนารวมทั้ง Android SDK (Software Development Kit) เตรียมไว้ให้กับนักพัฒนาได้เรียนรู้ และเมื่อนักพัฒนาต้องการจะเผยแพร่หรือจำหน่ายโปรแกรมที่พัฒนาแล้วเสร็จ แอนดรอยด์ก็ยังมีตลาดในการเผยแพร่โปรแกรมผ่าน Android Market โครงสร้างภาษาที่ใช้ในการพัฒนาสำหรับ Android SDK จะยึดโครงสร้างของภาษาจาวา (Java language) ในการเขียนโปรแกรม เพราะโปรแกรมที่พัฒนามาได้จะต้องทำงานอยู่ภายใต้ Dalvik Virtual Machine เช่นเดียวกับโปรแกรมจาวา ที่ต้องทำงานอยู่ภายใต้ Java Virtual Machine (Virtual Machine เปรียบได้กับสภาพแวดล้อมที่โปรแกรมทำงานอยู่) [3]

2.4 งานวิจัยที่เกี่ยวข้อง

ปัจจุบันได้มีงานวิจัยที่เกี่ยวข้องจำนวนมากที่นำเอาจีพีเอส กูเกิลแมพ เอพีไอ และระบบปฏิบัติการแอนดรอยด์บนสมาร์ตโฟน มาประยุกต์ใช้อย่างหลากหลาย โดยมีงานวิจัยที่ใช้เทคนิคที่คล้ายคลึงกับงานวิจัยนี้ ดังนี้

ศิริศักดิ์ [4] ได้ทำการวิจัย เรื่อง ระบบช่วยในการตัดสินใจเลือกโรงแรมโดยใช้เทคนิค GIS และ GPS บน Smart Phone งานวิจัยนี้ เป็นการพัฒนาระบบในรูปแบบของเว็บแอปพลิเคชัน (Web Application) ซึ่งเป็นการรวบรวมข้อมูลของโรงแรมเพื่อนำมาช่วยในการตัดสินใจเลือกเข้าพักโรงแรม เพื่อช่วยในการตัดสินใจและระบุระยะห่างระหว่างผู้ใช้กับโรงแรม มีเทคนิคที่ใช้คือ การกำหนดพิกัดโดยใช้ ระบบกำหนดตำแหน่งบนพื้นโลกผ่านดาวเทียม (GPS) ซึ่งจะทำการหาตำแหน่งของนักท่องเที่ยวและทำการเปรียบเทียบเทียบกับตำแหน่งของโรงแรมที่พักเพื่อหาโรงแรมที่ใกล้ที่สุดเพื่อช่วยนักท่องเที่ยวในการตัดสินใจเลือกโรงแรมที่พัก นำเสนอในรูปแบบแผนที่ เพื่อช่วยในการเดินทางไปยังโรงแรมที่พักที่ต้องการ ข้อเสนอแนะของงานวิจัยนี้คือ การพัฒนาระบบให้รองรับกิจกรรมที่เพิ่มขึ้น

ฉวีวรรณ [5] ได้ทำการวิจัยเรื่อง ระบบนำทางการท่องเที่ยวเชิงวัฒนธรรมผ่านดาวเทียมบอกพิกัด เป็นระบบบริการข้อมูลบนเว็บแอปพลิเคชันผ่านโทรศัพท์เคลื่อนที่ ด้วยการประยุกต์ใช้เทคโนโลยีกูเกิลแมพ (Google Maps) ซึ่งเป็นการบริการแผนที่นำมาระบุตำแหน่งของสถานที่ท่องเที่ยว ผสานกับเทคโนโลยีดาวเทียมบอกพิกัดแบบนำทาง (Global Positioning System:

GPS) บนโทรศัพท์เคลื่อนที่ ช่วยหาตำแหน่งของนักท่องเที่ยวและนำทางนักท่องเที่ยวให้สามารถไปยังสถานที่ท่องเที่ยวที่ต้องการได้อย่างถูกต้อง และหากตำแหน่งละติจูด ลองจิจูดของสถานที่ท่องเที่ยว และนักท่องเที่ยวใกล้เคียงกัน ระบบจะดึงข้อมูลจากฐานข้อมูลที่เก็บรายละเอียดต่าง ๆ ของสถานที่ท่องเที่ยว มาแสดงเป็นเสียงอธิบายประวัติความเป็นมาของสถานที่ท่องเที่ยว นั้น สิ่งที่น่าสนใจได้เสนอแนะไว้ คือ การพัฒนาเว็บแอปพลิเคชันที่สามารถรองรับโมบายเบราว์เซอร์ให้ได้ทุกตัว

Weifeng Shan, Jilin Feng, et al. [6] นำเสนองานวิจัยที่ใช้สมาร์ตโฟนในการเก็บรวบรวมข้อมูลภัยพิบัติแผ่นดินไหว โดยใช้สมาร์ตโฟนเก็บภาพ วิดีโอ และข้อมูลข้อความที่เป็นรายละเอียด ข้อมูลความต้องการของพื้นที่ที่เกิดภัยพิบัติแผ่นดินไหวร่วมกับการใช้จีพีเอส ในการระบุตำแหน่งผ่านการเปิดใช้งานอินเทอร์เน็ตบนสมาร์ตโฟน ผ่านไปยังเซิร์ฟเวอร์แสดงผลบนกูเกิลแมพ โดยใช้ไอคอนแตกต่างกัน เพื่อเลือกแสดงข้อมูล หรือ รูปภาพ วิดีโอ ทำให้การให้ความช่วยเหลือเป็นไปอย่างรวดเร็ว ในภาวะสถานการณ์ฉุกเฉิน โดยในอนาคตจำเป็นต้องปรับปรุงการทำงานและประสิทธิภาพของระบบเพื่อสนับสนุนการใช้งานใช้กับภัยพิบัติอื่นมากขึ้น เช่น น้ำท่วม พายุไต้ฝุ่น ไฟไหม้ อุบัติเหตุทางจราจร และภัยอื่น ๆ และการวิเคราะห์เชิงพื้นที่ ผ่านสมาร์ตโฟนมากขึ้น

Yu-Chih Liao, Chen-Chia Chuang, et al. [7] นำเสนองานวิจัยที่ออกแบบระบบที่ระบุตำแหน่งจีพีเอส ที่อยู่บนสมาร์ตโฟนระบบปฏิบัติการแอนดรอยด์ ที่สามารถระบุตำแหน่งปัจจุบัน พร้อมละติจูด ลองจิจูดและที่อยู่ แสดงบนกูเกิลแมพ ที่สามารถเพิ่มเบอร์ติดต่อฉุกเฉิน อีเมล และข้อความอื่นๆ ทำการถ่ายรูปหน้าจอที่แสดงตำแหน่งพร้อมข้อความการติดต่อทั้งหมด เพื่อใช้สำหรับส่ง ข้อความหรืออีเมล ในกรณีฉุกเฉิน โดยสิ่งที่ต้องพัฒนาและปรับปรุงต่อไป คือการเพิ่มชั้นของแผนที่รวมถึงเพิ่มตำแหน่งของบริเวณที่ใกล้เคียงได้ เพื่อใช้งานร่วมกันได้มากขึ้น

จากการศึกษางานวิจัยที่เกี่ยวข้อง ได้มีการนำเทคนิคจีพีเอส มาใช้ในการระบุตำแหน่งบนแผนที่กูเกิลแมพ และการพัฒนาแอปพลิเคชันบนสมาร์ตโฟนระบบปฏิบัติการแอนดรอยด์ ผู้วิจัยจึงมีแนวคิดในการพัฒนาระบบติดตามสถานการณ์

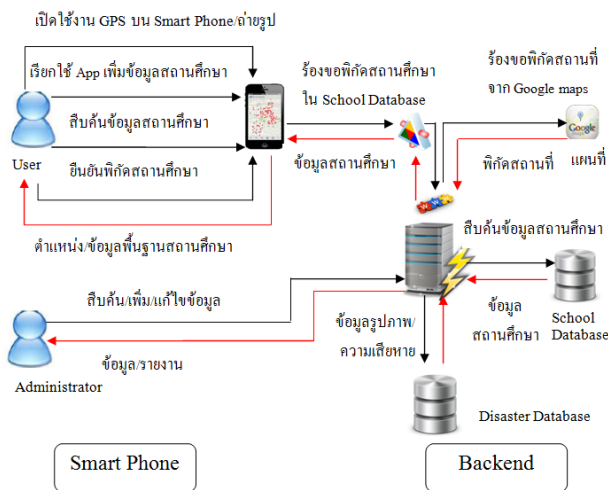
สถานศึกษาที่ประสบอุทกภัย โดยใช้พิกัดจีพีเอสของสถานศึกษาจากฐานข้อมูล ที่สามารถแสดงตำแหน่งปัจจุบันพร้อมข้อมูลสถานศึกษา และให้สามารถส่งข้อมูลรูปถ่ายเพื่อความรวดเร็ว ถูกต้อง และแม่นยำในการรายงานสถานการณ์

3. วิธีการดำเนินการวิจัย

จากแนวความคิดในการพัฒนาระบบติดตามสถานการณ์สถานศึกษาที่ประสบอุทกภัยด้วยสมาร์ตโฟนระบบปฏิบัติการแอนดรอยด์ เพื่อให้การรายงานและติดตามสถานการณ์ได้อย่างรวดเร็ว นั้น จึงออกแบบระบบที่จะนำเอาพิกัดจีพีเอสของสถานศึกษาที่มีอยู่ในฐานข้อมูลไปแสดงบนกูเกิลแมพ ที่พัฒนาบนสมาร์ตโฟนระบบปฏิบัติการแอนดรอยด์ และให้สามารถอัปโหลดรูปถ่าย พร้อมระบุระดับความเสียหาย และระดับน้ำ สามารถเพิ่มพิกัดสถานศึกษาที่ต้องการจากพิกัดจีพีเอสของรูปถ่ายได้ในกรณีที่สถานศึกษาบางแห่งไม่มีพิกัดในฐานข้อมูล โดยมีวิธีการดำเนินการ ดังนี้

3.1 การออกแบบระบบ

การออกแบบระบบแบ่งเป็น 2 ส่วน คือ ผู้ใช้งานบนสมาร์ตโฟน และผู้ดูแลระบบ โดยการออกแบบสถาปัตยกรรมระบบ แสดงไว้ดังภาพที่ 1 ดังนี้



ภาพที่ 1: การออกแบบสถาปัตยกรรมระบบ

ส่วนที่ 1 ผู้ใช้งานบนสมาร์ตโฟน เปิดใช้งานแอปพลิเคชันพร้อมการเปิดใช้งานจีพีเอส ให้ระบบแสดงตำแหน่งปัจจุบันของผู้ใช้งานบนกูเกิลแมพ ผู้ใช้งานสามารถยืนยันความถูกต้องของตำแหน่งสถานศึกษา สามารถค้นหาจากชื่อของสถานศึกษา

เพื่อให้กูเกิลแมพแสดงตำแหน่งที่ต้องการ และรายงานข้อมูลเข้าระบบ สามารถเรียกดูข้อมูลที่รายงานได้

ส่วนที่ 2 Backend บนเว็บเซิร์ฟเวอร์ เมื่อผู้ใช้เปิดใช้งานแอปพลิเคชัน จะทำการร้องขอพิกัดของสถานศึกษาผ่าน Web Services เพื่อทำการสืบค้นจากฐานข้อมูล School Database และร้องขอพิกัดสถานที่จากกูเกิลแมพ ส่งกลับไปยังผู้ใช้งานเพื่อแสดงผล สำหรับข้อมูลการประสพภัยและรูปถ่ายที่ผู้ใช้งานอัปโหลดเข้าระบบจะถูกเก็บไว้ที่ Disaster Database

3.2 การพัฒนาระบบ

พัฒนาระบบโดยใช้ Android SDK บนโทรศัพท์มือถือระบบปฏิบัติการ Android 4.2 ขึ้นไปเพื่อให้รองรับ Google Map API version 2 ใช้ Eclipse Development Tools and Java Development Kit (JDK) ในส่วนของ Web-based Application ใช้ CMS และ Google Map API version 3 เชื่อมโยงกับฐานข้อมูลสถานศึกษาของกระทรวงศึกษาธิการ โดยใช้เว็บเซอร์วิส

3.3 การทดสอบประสิทธิภาพของระบบ

กำหนดวิธีทดสอบใช้ Black Box Testing ทดสอบการทำงานบนสมาร์ตโฟน และ Backend โดยผู้พัฒนาระบบเป็นผู้ดำเนินการทดสอบ จากตำแหน่งของสถานศึกษา จำนวน 30 แห่ง ดังนี้

3.3.1 ทดสอบการทำงานบนสมาร์ตโฟน

เมื่อเปิดใช้งานระบบ สามารถแสดงตำแหน่งปัจจุบันและชื่อสถานศึกษาบนแผนที่ สามารถส่งรูปถ่ายหรือเลือกรูปถ่ายจากหน่วยความจำของโทรศัพท์ ระบุระดับความเสียหายระดับน้ำ ส่งข้อมูลเข้าระบบได้ และเมื่อเรียกดูสามารถแสดงข้อมูลที่รายงานผ่านกูเกิลแมพ ได้ถูกต้องทั้งหมด

ทดสอบเพิ่มพิกัดของสถานศึกษาใหม่ หากสถานศึกษาบางแห่งไม่มีพิกัดในฐานข้อมูล โดยใช้พิกัดจีพีเอสจากรูปถ่ายในการระบุตำแหน่ง และส่งข้อมูลเข้าระบบได้

ทดสอบปรับปรุง แก้ไขตำแหน่งของสถานศึกษาเดิม หากตำแหน่งของสถานศึกษาที่แสดงมีความคลาดเคลื่อน โดยใช้พิกัดจีพีเอส จากรูปถ่ายในการระบุตำแหน่ง และส่งข้อมูลเข้าระบบได้

3.3.2 ในส่วน Backend Server

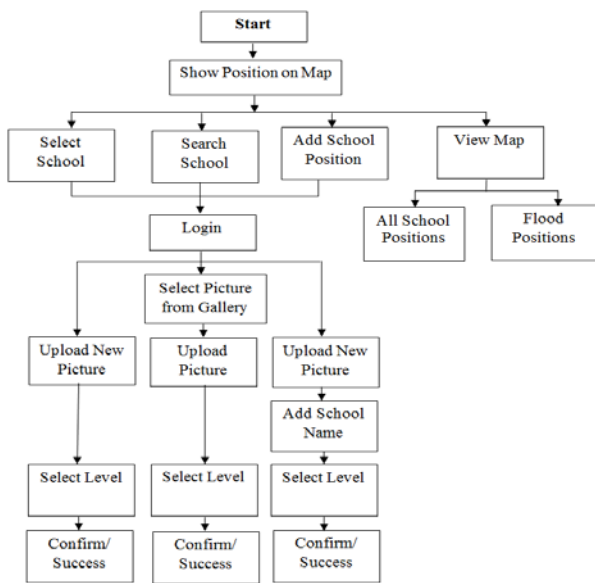
ทดสอบการเรียกรายงาน สถานศึกษาที่ประสบอุทกภัย

สถานศึกษาที่แจ้งเพิ่มพิกัดใหม่ และปรับปรุงแก้ไขตำแหน่งสถานศึกษาเดิม ปรากฏรายชื่อสถานศึกษาที่รายงาน ครบถ้วนตามจำนวนที่ได้มีการทดสอบส่งข้อมูลผ่านสมาร์ตโฟน และทดสอบส่งออกข้อมูลการรายงานในรูปแบบไฟล์ Excel ได้

4. ผลการดำเนินงาน

4.1 ผลการพัฒนาระบบ

ผลการพัฒนาระบบได้ดำเนินการตามกระบวนการและขั้นตอนที่ได้ออกแบบไว้ โดยได้ทำการติดตั้งระบบที่พัฒนาบนสมาร์ตโฟนระบบปฏิบัติการแอนดรอยด์ และมีกระบวนการทำงาน ดังภาพที่ 2



ภาพที่ 2: กระบวนการทำงานระบบ

บนสมาร์ตโฟน ผู้ใช้งานทั่วไปเปิดใช้งานแอปพลิเคชันพร้อมการเปิดใช้งานจีพีเอส สามารถค้นหาตำแหน่งของสถานศึกษาจากชื่อ ค้นหาโดยเลือกแสดงเฉพาะสถานศึกษาที่ประสบภัย หรือเลือกตำแหน่งของสถานศึกษาทั้งหมด และเรียกดูข้อมูลการรายงานผ่านกูเกิลแมปได้

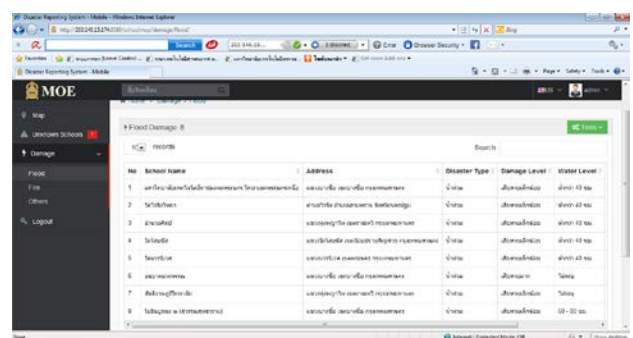
ผู้ใช้งานที่ต้องการรายงานสถานการณ์ สามารถเลือกตำแหน่งของสถานศึกษาที่เป็นตำแหน่งปัจจุบันหรือค้นหาจากชื่อ และ Login ก่อนการรายงาน อัปโหลดรูปถ่าย พร้อมระบุระดับความเสียหาย ระดับน้ำ เพื่อบันทึกข้อมูลเข้าระบบ กรณีที่ระบบไม่แสดงตำแหน่งปัจจุบันหรือค้นหาไม่พบชื่อสถานศึกษานั้น ให้ผู้รายงานกรอกชื่อสถานศึกษา และรายงานสถานการณ์ได้ โดยหน้าจอระบบบนสมาร์ตโฟนในขั้นตอนการรายงาน

แสดง ข้อมูล ของ สถาน ศึกษา ละ ดิจิต ลอจ จิอุค และรายละเอียดของการประสบภัยพร้อมรูปถ่าย และเมื่อการรายงานสมบูรณ์ แสดงตำแหน่งพร้อมข้อมูลบนกูเกิลแมป ดังภาพที่ 3

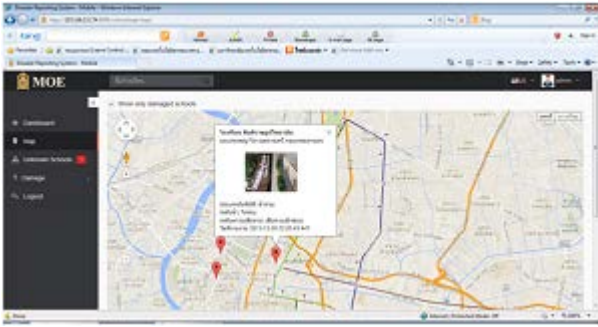


ภาพที่ 3: หน้าจอแสดงตำแหน่งและข้อมูลสถานศึกษาบนสมาร์ตโฟน

ผู้ดูแลระบบ ในส่วนของ Backend สามารถเรียก รายงาน ประกอบด้วย แผนที่ รายชื่อสถานศึกษาที่ประสบภัย และรายชื่อสถานศึกษาที่ประสบภัยแต่ยังไม่มียังข้อมูลพื้นฐานของสถานศึกษานั้น โดยระบบจะแสดงจำนวนสถานศึกษาที่รอการตรวจสอบ (Unknown) ผู้ดูแลระบบตรวจสอบและปรับปรุงข้อมูลโดยใช้รหัสสถานศึกษาเชื่อมโยงกับข้อมูลพื้นฐานของสถานศึกษา โดยรายชื่อสถานศึกษาที่ปรับปรุงแล้วจะถูกเปลี่ยนสถานะไปอยู่ในรายชื่อสถานศึกษาที่ประสบภัย และสามารถส่งออกรายงานเป็นไฟล์ Excel ได้ และเมื่อปรับปรุงข้อมูลแล้วสามารถเรียกดูและแสดงตำแหน่งพร้อมข้อมูลบนกูเกิลแมป โดยรูปแบบรายงานและหน้าจอแสดงแผนที่ ดังภาพที่ 4 และภาพที่ 5



ภาพที่ 4: หน้าจอรายงาน



ภาพที่ 5: หน้าจอแสดงตำแหน่งและข้อมูลการรายงาน

4.2 ผลการประเมินประสิทธิภาพระบบ

ผลการประเมินประสิทธิภาพของระบบ โดยทำการทดสอบตำแหน่งของสถานศึกษา จำนวน 30 แห่ง ให้แสดงตำแหน่งปัจจุบัน และระบุ ชื่อ ที่ตั้งของสถานศึกษา ข้อมูลระดับความเสียหาย ระดับน้ำ และรูปถ่าย เรียกดูข้อมูลที่รายงานบนกูเกิลแมพ และในส่วน Web-Based Application ให้แสดงรายชื่อสถานศึกษา เพิ่ม/แก้ไขข้อมูล และแสดงผลผ่านกูเกิลแมพ ระบบสามารถแสดงผลได้อย่างถูกต้องทั้งหมด

5. สรุปผลและข้อเสนอแนะ

การพัฒนาบบติดตามสถานการณ์สถานศึกษาที่ประสบอุทกภัยด้วยสมาร์ตโฟนระบบปฏิบัติการแอนดรอยด์ เพื่อใช้ในการรายงานสถานการณ์ให้มีความสะดวกรวดเร็ว จากเดิมผู้รายงานต้องกรอกข้อมูลในแบบฟอร์มของหน่วยงาน ทำให้มีความล่าช้า และในปัจจุบันสมาร์ตโฟนมีผู้นิยมใช้งานจำนวนมาก รวมถึงระบบปฏิบัติการแอนดรอยด์ เป็น Open Source ที่สามารถพัฒนาแอปพลิเคชันต่าง ๆ ได้อย่างกว้างขวาง

การพัฒนาบบใช้ Android SDK เชื่อมโยงกับฐานข้อมูลสถานศึกษา พัฒนบบสมาร์ตโฟนระบบปฏิบัติการแอนดรอยด์ การทำงานแบ่งออกเป็น 2 ส่วน คือ การทำงานบนสมาร์ตโฟน ประกอบด้วย การเลือกตำแหน่งปัจจุบัน การค้นหาสถานศึกษา การเพิ่มพิกัดสถานศึกษา การแสดงตำแหน่งสถานศึกษา บนแผนที่ การทำงานในส่วน Backend Server เป็นส่วนที่ผู้ดูแลระบบสามารถเรียกรายงาน ปรับปรุงแก้ไขข้อมูล และส่งออกรายงานเป็นไฟล์ Excel ได้

จากผลการประเมินประสิทธิภาพของระบบ สามารถระบุตำแหน่งปัจจุบัน ชื่อ ที่ตั้งสถานศึกษาได้อย่างถูกต้อง สามารถอัปโหลดรูปและข้อมูล การค้นหารายชื่อสถานศึกษา เพิ่มพิกัด

สถานศึกษาโดยใช้พิกัดจีพีเอสจากรูปถ่าย จำนวน 30 แห่ง สามารถแสดงผลได้อย่างถูกต้องทั้งหมด สรุปได้ว่าระบบมีความถูกต้องเท่ากับ 100 เปอร์เซ็นต์

ปัญหาและอุปสรรคที่พบคือ ความเร็วของเครือข่ายของสัญญาณอินเทอร์เน็ต มีผลต่อการค้นหาตำแหน่งปัจจุบันและการอัปโหลดรูป โดยแนะนำว่าควรเป็นเครือข่าย 3G

ข้อเสนอแนะจากงานวิจัย การพัฒนาระบบติดตามสถานการณ์สถานศึกษาที่ประสบอุทกภัยโดยใช้สมาร์ตโฟนระบบปฏิบัติการแอนดรอยด์ ในอนาคตควรมีการพัฒนาบนระบบปฏิบัติการอื่น เพื่อให้สามารถใช้งานได้ทั้งถึงและควรเพิ่มระบบการแจ้งเตือนด้วยและอาจจะมีการติดตามรายงานเพื่อให้สถานศึกษาทราบถึงการช่วยเหลือได้

เอกสารอ้างอิง

- [1] Global5 Co., Ltd. “ความรู้ทั่วไปเกี่ยวกับ GPS” [ออนไลน์]. เข้าถึงได้จาก <http://www.global5thailand.com/thai/gps.htm#1>. GPS_คืออะไร.สืบค้น 1 กันยายน 2556.
- [2] “รู้จักกับGoogle Map API” [ออนไลน์]. เข้าถึงได้จาก <http://mapsmile.com/mapboard/google-maps-api/google-map-api/ msg8/?PHPSESSID=cbd794c28489d3e2d2050662fb1772d6#msg8>.สืบค้น.1กันยายน 2556.
- [3] 9M / webmaster@sourcecode.in.th. “รู้จักกับแอนดรอยด์” [ออนไลน์]. เข้าถึงได้จาก <http://www.sourcecode.in.th/articles.php?id=71>.สืบค้น.1กันยายน 2556.
- [4] ชีรศักดิ์ อธิภาคย์, ระบบช่วยในการตัดสินใจเลือกโรงแรมโดยใช้เทคนิค GIS และ GPS บน Smart Phone, คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2554.
- [5] จวีวรรณ พุ่งสิบสาม, ระบบนำทางการท่องเที่ยวเชิงวัฒนธรรมผ่านดาวเทียมบอกพิกัด, สาขาวิชาวิศวกรรมซอฟต์แวร์ มหาวิทยาลัยเชียงใหม่, 2554.
- [6] Weifeng Shan, Jilin Feng, et al., “Collecting Earthquake Disaster Area Information Using Smart Phone” 2012 *International Conference on System Science and Engineering*, June 30-July 2, 2012, Dalian, China, pp.310-314, 2012.
- [7] Yu-Chih Liao, Chen-Chia Chuang, et al., “Systematic Design for the Global Positional Systems with Application in Intelligent Google Android Phone” *Proceedings of 2011 International Conference on System Science and Engineering*, Macau, China, June 2011, pp.182-186, 2011.

ระบบมัลติเทคที่พกพาสำหรับการท่องเที่ยว Portable Guide System for Tourism

ขวัญชัย สุขแสน(Khwanchai Suksaen)¹ ณัฐวุฒิ มูลศาสตร์(Nattawut Moonsart)²และ มนัสนันท์ จันทโท(Manassanant chantho)³

สาขาวิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏอุตรธานี

khwanchais@gmail.com¹, nattawut.moonsart@gmail.com², manassanantchantho@gmail.com³

บทคัดย่อ

ข้อมูลการท่องเที่ยวส่วนใหญ่จะนำเสนอเป็นหนังสือหรือเว็บไซต์ แต่อย่างไรก็ดี นักท่องเที่ยวที่นิยมเดินทางท่องเที่ยวด้วยตนเองมักต้องการข้อมูลหรือคำแนะนำเกี่ยวกับแหล่งท่องเที่ยวที่จะเดินทางไปอย่างทันที ดังนั้น มัลติเทคที่พกพาสำหรับการท่องเที่ยวจึงถูกพัฒนาขึ้นมาเป็นแอปพลิเคชันสำหรับแนะนำและให้สารสนเทศเกี่ยวกับสถานที่ท่องเที่ยว โรงแรม และ ร้านอาหารจากตำแหน่งพิกัดภูมิศาสตร์ของสถานที่เหล่านั้น ซึ่งแอปพลิเคชันนี้ทำงานบนระบบปฏิบัติการแอนดรอยด์ ที่ผู้ใช้ที่เคยไปท่องเที่ยวยังสถานที่ต่างๆ สามารถทำการแนะนำสถานที่แห่งใหม่เข้าสู่ระบบ หลังจากนั้นต้องผ่านการตรวจสอบความเหมาะสมจากผู้ดูแลระบบก่อนที่สถานที่แห่งนั้นจะเสนอเป็นสถานที่แนะนำต่อสาธารณชนได้ สำหรับกระบวนการแนะนำแหล่งท่องเที่ยวนั้นจะใช้เทคนิคการหาสมาชิกที่ใกล้ที่สุด (*k*-Nearest Neighbor) จากข้อมูลเกี่ยวกับข้อคิดเห็นและคะแนนความนิยมด้านความสวยงาม ด้านความคุ้มค่า ด้านการให้บริการ ด้านความสะดวก และด้านเส้นทางการเดินทาง ที่ผู้ท่องเที่ยวคนอื่นๆ ได้ให้คะแนนความนิยมไว้เทียบกับความนิยมของผู้ใช้ที่กำลังใช้งานอยู่ จากการประเมินความพึงพอใจของแอปพลิเคชันโดยกลุ่มตัวอย่างผู้ใช้งานทั่วไป มีความพึงพอใจอยู่ในระดับดี

คำสำคัญ: มัลติเทค, ระบบให้คำแนะนำ, เทคนิคเพื่อนบ้านที่ใกล้ที่สุด, ตำแหน่งพิกัดทางภูมิศาสตร์, แหล่งท่องเที่ยว

Abstract

Tourist information mostly is presented in books or websites. However, popular tourists traveling by themselves often need information or advice about tourist

destinations to travel immediately. Therefore, the portable guide system for traveling has been developed as an application for advice and present information about tourist attractions, hotels and restaurants of geographic coordinates of those places. The developed application works on Android operating system. The users who travel to different places can advise new places into the system, and then the system administrator have to verify the suitability of a system prior to present such places publicly. For the procedure of tourist recommendation, it uses the *k*-Nearest Neighbor technique based on the information about comments and popular ratings about view point of beauty, worthiness, service, cleanness, and travel route from other travelers compared to the popularity rating from the active user. The satisfy evaluation of this application by end users is resulted in good level.

Keywords: Guide, Recommendation System, *k*-Nearest Neighbor, Geographic coordinates, Tourist attraction

1. บทนำ

นักท่องเที่ยวนิยมเดินทางท่องเที่ยวด้วยตนเอง อาจมีความต้องการที่จะได้รับข้อมูลหรือคำแนะนำของแหล่งท่องเที่ยว ณ จุดที่ยืนอยู่ หรืออ่านคำแนะนำจากผู้ที่เคยมาเที่ยวก่อนหน้านี้ เพื่อที่จะแนะนำสถานที่ท่องเที่ยวได้ และการนำเสนอข้อมูลที่เกี่ยวข้องกับการท่องเที่ยว รวมถึงการพัฒนาด้านการบริการ ข้อมูลการท่องเที่ยว มักจะเป็นการให้บริการข้อมูลในรูปแบบเว็บไซต์ ในทางตรงกันข้ามการบริการข้อมูล ก็มีข้อมูลที่แตกต่างหรือมีข้อมูลที่ไม่ค่อยมีการเปลี่ยนแปลงสม่ำเสมอ และ

สืบเนื่องจากการบริการข้อมูลการท่องเที่ยว ที่มีบริการด้านข้อมูล รูปภาพ และเส้นทางการเดินทาง ซึ่งอาจจะยังไม่เป็นปัจจุบัน ตามที่นักท่องเที่ยวต้องการ ทางผู้พัฒนามีแนวคิดที่จะพัฒนาโมบาย แอปพลิเคชัน สำหรับบริการข้อมูลการท่องเที่ยว แนะนำสถานที่ท่องเที่ยว พร้อมทั้งยังสามารถแนะนำโรงแรม และร้านอาหาร โดยให้นักท่องเที่ยวกรอกความต้องการหรืองบประมาณ ลงในแอปพลิเคชัน เพื่อให้แอปพลิเคชันแสดงผลข้อมูลที่ตรงกับความต้องการนั้นๆ ทำให้สะดวกรวดเร็วในการวางแผนการเดินทาง และไม่ต้องเสียค่าใช้จ่ายในการจ้างมัคคุเทศก์

ดังนั้น ทางผู้พัฒนาได้พัฒนาระบบงานบนโทรศัพท์เคลื่อนที่ที่ใช้ระบบปฏิบัติการแอนดรอยด์ ที่ช่วยแนะนำสถานที่ท่องเที่ยว โรงแรม ร้านอาหาร ที่น่าสนใจ โดยได้แนะนำสถานที่ต่างๆ ให้ตรงตามปัจจัยความต้องการของผู้ใช้งาน โดยที่ระบบมีข้อมูลการท่องเที่ยวมีการเปลี่ยนแปลงสม่ำเสมอ นักท่องเที่ยวสามารถเข้าใจได้ง่าย เป็นสื่อกลางที่ช่วยแนะนำสถานที่ท่องเที่ยว การค้นหาสถานที่ท่องเที่ยว โรงแรม ร้านอาหาร และเส้นทางการเดินทางได้

2. ระบบการให้คำแนะนำ

ระบบให้คำแนะนำ (Recommendation System) เป็นระบบที่ถูกนำมาใช้เพื่อพยายามแนะนำเสนอสินค้าบริการและการแนะนำการเลือกสิ่งต่างๆ ที่คาดว่าผู้ใช้น่าจะสนใจหรืออาจจะเป็นข้อมูลที่ใช้ต้องการ ซึ่งระบบให้คำแนะนำจะประกอบไปด้วย 4 ส่วนหลักๆ[1] ได้แก่ (1) ส่วนข้อมูลพื้นฐานที่จำเป็นต้องใช้ในการประมวลผล เช่น โปรไฟล์ของผู้ใช้งานแต่ละคน (2) ส่วนการป้อนข้อมูล เป็นข้อมูลที่ได้จากการป้อนข้อมูลเข้ามาของผู้ใช้ เช่น การให้คะแนนระดับความสำคัญ ซึ่งมีอยู่ 2 แบบคือแบบชัดเจน (Explicit) เช่น ระดับความนิยมตั้งแต่ 1 ถึง 5 และ แบบไม่ชัดเจน (Implicit) เช่น ประวัติการใช้งานของระบบในอดีต (3) ส่วนของอัลกอริทึมเป็นส่วนสำคัญที่สุดที่ใช้ประมวลผลข้อมูลเพื่อให้การแนะนำข้อมูลออกมาให้ตรงกับความต้องการของผู้ใช้มากที่สุด และ (4) ส่วนของการนำเสนอคำแนะนำ หรือนำเสนอสินค้าที่ตรงกับความต้องการของผู้ใช้มากที่สุด

3. คุณลักษณะของข้อมูลของสถานที่ท่องเที่ยว

เป็นวิธีการหนึ่งของการทำ การกรองข้อมูล โดยอาศัยเนื้อหาเป็นหลัก โดยวิธีการทำงานของเทคนิคนี้คือ จะให้ความสนใจเนื้อหาของข้อมูลเป็นสำคัญ เช่น คุณลักษณะ (Feature) เพื่อค้นหาข้อมูลที่ผู้ใช้นั้นสนใจ ซึ่งวิธีการของเทคนิคดังกล่าวจะไม่ประสบกับปัญหาการให้ระดับความสำคัญ ต่อชั้นข้อมูลที่ไม่ทั่วถึง และปัญหาชั้นข้อมูลที่ขังไม่ได้ให้ระดับความสำคัญ

ดังนั้นวิธีการนี้จะต้องมีการคำนวณหาค่าความคล้ายคลึงกัน (Similarity measure) ระหว่างข้อมูลหรือสินค้าที่ระบบมีอยู่กับความ ต้องการในตัวข้อมูลหรือสินค้าที่ถูกค่าต้องการ ซึ่งอัลกอริทึมที่นิยมใช้มากที่สุดคือ วิธีการหาสมาชิกที่ใกล้กันที่สุด k-NN (k-Nearest Neighbor) ซึ่งจะต้องมีการคำนวณหาค่าระยะห่างระหว่างข้อมูล ข้อดีคือให้ผลลัพธ์ที่มีความถูกต้องสูง เรียนรู้เร็ว และรองรับข้อมูลจำนวนมาก [2][3]

3.1 สร้างเมตริกซ์ของชั้นข้อมูลและเมตริกซ์ความต้องการของผู้ใช้

ขั้นตอนแรกทำการสร้างเมตริกซ์ของชั้นข้อมูล โดยมีรูปแบบเป็นตารางที่มีแถวเป็นสถานที่ท่องเที่ยวและมีคอลัมน์เป็นปัจจัยในด้านต่างๆ ของสถานที่ท่องเที่ยว แต่ละสถานที่

ตารางที่ 1: เมตริกซ์ของชั้นข้อมูล

Feature ลำดับที่ \	ชื่อสถานที่ท่องเที่ยว	จังหวัด	ด้านความสวยงาม	ด้านความสะอาด	ด้านสถานที่ตั้ง	...	งบประมาณ
1	ศาลเจ้าปู่-ย่า	อุดรธานี	4	4	4	...	150
2	ศาลหลักเมือง	อุดรธานี	4	4	4	...	50
3	ภูผอยลม	อุดรธานี	5	3	3	...	500
4	ทะเลบัวแดง	อุดรธานี	4	3	3	...	250
5	วัดป่าบ้านตาด	อุดรธานี	4	4	3	...	50

ขั้นตอนที่สองทำการสร้างเมตริกซ์ความต้องการของผู้ใช้ โดยจะอยู่ในรูปแบบของพีเชอร์เวกเตอร์ (Feature vector) ซึ่งข้อมูลในส่วนของความต้องการของผู้ใช้นั้นจะประกอบไปด้วย 2 ส่วน คือ ส่วนแรกเป็นรายละเอียดของสถานที่ท่องเที่ยวที่ผู้ใช้ได้ทำการระบุความต้องการ (Request) และส่วนที่สองเป็นค่าน้ำหนักความสำคัญของแต่ละส่วนประกอบ (Weight) ซึ่งจะมีระดับความสำคัญ ตั้งแต่ 1-5

ตารางที่ 2:เมตริกซ์ความต้องการของผู้ใช้

Feature	จังหวัด	ด้านความ สวยงาม	ด้านความ สะอาด	ด้านสถาน ที่ที่ตั้ง	...	งบประมาณ	สถานที่ ท่องเที่ยว
request	อุดรธานี	-	-	-	...	150	?
weight	-	4	4	3	...	-	

3.2คำนวณหาค่าความคล้ายคลึงกัน (Similarity Measure)

ขั้นตอนของการคำนวณทั้ง 2 พารามิเตอร์เพื่อทำการแบ่งคลาส โดยใช้เทคนิคของClassificationซึ่งอัลกอริทึมที่ใช้คือ k-NN (k-Nearest Neighbor)โดยมีวิธีการคำนวณคือ เมื่อมีข้อมูลใหม่เข้ามาในระบบจะดึงข้อมูลอื่นๆ ที่มีอยู่และมีลักษณะคล้ายคลึงกันที่สุด kตัวออกมา

รูปแบบของข้อมูล ข้อมูลที่นำมาทำการคำนวณประกอบไปด้วย n แอททริบิวต์ โดยข้อมูลในเรคคอร์ด Xใดๆ แทนด้วยเวกเตอร์ได้ $X = \langle X_1, X_2, X_3, \dots, X_n \rangle$ โดย X_i หมายถึงค่าข้อมูลในแอททริบิวต์ที่ i เรคคอร์ด X ใดๆ กำหนดค่าระดับความสำคัญ $C = 5$ การคำนวณระยะทาง ในการคำนวณระยะทาง $d(X, Y)$ หมายถึงระยะห่างระหว่างข้อมูล X และ Y โดยมีการประยุกต์ใช้การวัดระยะทางของ Standard Euclidean Distance กับการให้น้ำหนักหรือการระดับให้มีความสำคัญกับแอททริบิวต์แต่ละตัวซึ่งเรียกว่า Distance Weighted เพื่อให้เหมาะกับการคำนวณหาค่าความคล้ายคลึงกันของข้อมูลสถานที่ท่องเที่ยวดังแสดงในสมการ

$$d(x, y) = \sqrt{\sum_{i=1}^n (c - w_i)(x_i - y_i)^2} \tag{1}$$

โดยที่

C คือ ค่าระดับการความสำคัญ คือ 1-5

X คือ จำนวนของข้อกำหนดของสถานที่ท่องเที่ยวที่ผู้ใช้ต้องการ

Y คือ จำนวนของข้อกำหนดของสถานที่ท่องเที่ยวที่มีอยู่ในระบบ

W คือ ค่าของน้ำหนักของข้อมูลแต่ละแอททริบิวต์

x_i คือ ค่าข้อมูลในแอททริบิวต์ที่ i ของเรคคอร์ด X

y_i คือ ค่าข้อมูลในแอททริบิวต์ที่ i ของเรคคอร์ด Y

w_i คือ ค่าน้ำหนักของข้อมูลในแอททริบิวต์ที่ i

4.งานวิจัยที่เกี่ยวข้อง

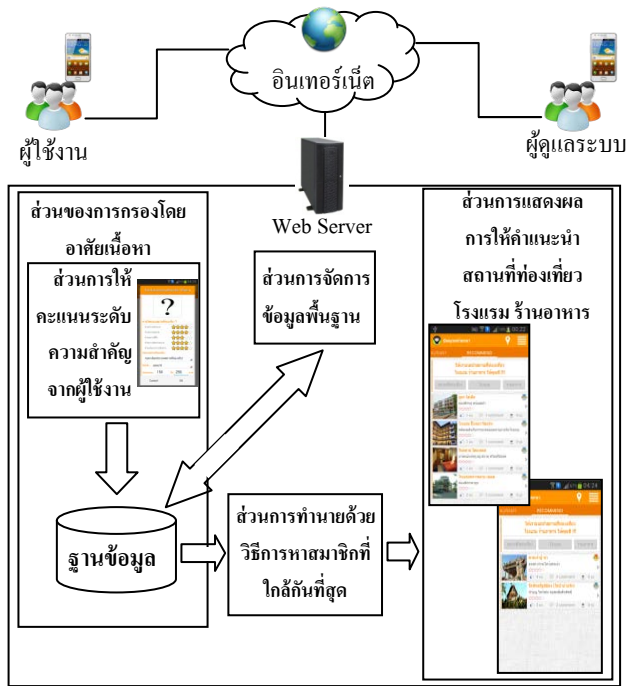
เกวลินและณัฐวิ[4] ได้พัฒนาระบบช่วยแนะนำในการเลือกซื้อเครื่องประดับผ่านเครือข่ายอินเทอร์เน็ตกรณศึกษา ร้านที่นแกลลอรี่ที่มีความเหมาะสมกับผู้ใช้แต่ละบุคคลโดยใช้เทคนิคการจัดกลุ่ม(Cluster)ในการจัดกลุ่มข้อมูลและคุณสมบัติของเครื่องประดับ ผลการทำงานของระบบมีประสิทธิภาพอยู่ในระดับดี และสามารถนำไปใช้งานได้อย่างมีประสิทธิภาพ

ศุภรัตน์ [5] ได้พัฒนาระบบให้คำแนะนำในการเลือกซื้อเลนส์สำหรับกล้องดีเอสแอลอาร์ โดยการใช้เทคนิค Content-based Filteringร่วมกับเทคนิค Collaborative Filtering เทคนิค Content-based Filteringทำการแนะนำเลนส์กล้องจากข้อมูลในระบบ ซึ่งมีอยู่ทั้งหมด 283 เลนส์ ที่มีความคล้ายคลึงกับคุณลักษณะของเลนส์กล้องที่ผู้ใช้ต้องการ สำหรับเทคนิค Collaborative Filteringระบบทำการแนะนำข้อมูลจากชิ้นข้อมูลในระบบ เทียบกับเปลี่ยนค่าว่าระดับความสำคัญของผู้ใช้งานและทำนายเลนส์กล้องที่คาดว่าตรงกับความต้องการของผู้ใช้

จากการศึกษาทฤษฎี บทความและงานวิจัยที่เกี่ยวข้องข้างต้นพบว่าระบบให้คำแนะนำข้อมูลเป็นระบบที่ได้รับความนิยมและมีประสิทธิภาพมากในการนำมาประยุกต์ใช้แนะนำสินค้าและบริการ เพื่อตอบสนองความต้องการของผู้ใช้ ซึ่งเลือกเทคนิค Content-based filteringสำหรับนำมาพัฒนาระบบนี้ ดังนั้นงานวิจัยนี้ได้นำเอาเทคนิคดังกล่าวมาประยุกต์ใช้เพื่อสร้างระบบแนะนำสถานที่ท่องเที่ยว โรงแรมและร้านอาหารซึ่งจะพิจารณาจากคุณลักษณะคุณลักษณะพื้นฐานต่างๆ ของสถานที่ที่ผู้ใช้สนใจเป็นสำคัญ[6][7]

5.แนวทางการดำเนินงาน

ในการพัฒนาโปรแกรมมีคฤเทศก์พทพบบระบบปฏิบัติการแอนดรอยด์ จากข้อมูลที่ได้รวบรวมมานั้นสามารถสรุปองค์ประกอบที่เกี่ยวข้องและ โมดูลการทำงานหลักของระบบได้ดังภาพที่ 1



ภาพที่ 1:ระบบ การให้คำแนะนำ กับผู้ใช้งาน

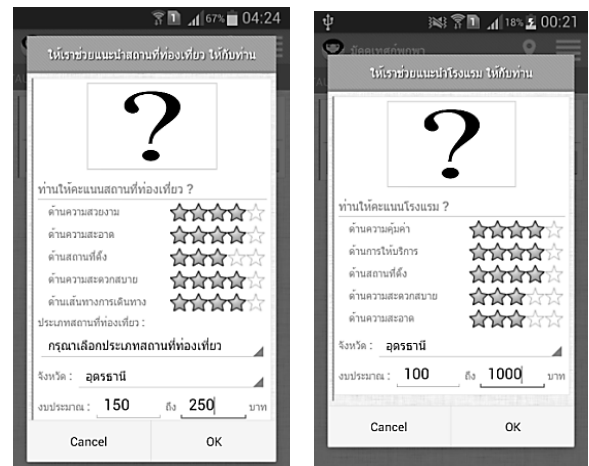
ผู้ใช้งานของระบบคือผู้ดูแลระบบและผู้ใช้งานทั่วไป ผ่านทางแอปพลิเคชันร่วมกับสมาร์ตโฟน บนระบบปฏิบัติการแอนดรอยด์ ส่วนการทำงานทั้งหมดประมวลผลทำงานบนเว็บเซิร์ฟเวอร์ มีการร้องขอและตอบสนองส่งผ่านข้อมูลระหว่างกันผ่านเครือข่ายอินเทอร์เน็ต ส่วนการทำงานหลักประกอบด้วย (1)ส่วนการจัดการข้อมูลพื้นฐาน ซึ่งจะเกี่ยวข้องกับข้อมูลแหล่งท่องเที่ยว โรงแรม ร้านอาหาร มีรายละเอียดเกี่ยวกับประเภทที่อยู่ การติดต่อ พิกัดภูมิศาสตร์ (2) ส่วนการให้คะแนนระดับความสำคัญจากผู้ใช้งาน โดยผู้ใช้งานจะต้องบันทึกความเห็นโดยให้คะแนนระดับความสำคัญ ในแต่ละด้าน (3) ส่วนการทำนายด้วยวิธีการหาสมาชิกที่ใกล้เคียงที่สุด เป็นนำข้อมูลการให้ระดับความสำคัญจากผู้ใช้งาน เพื่อนำไปประมวลผลหาค่าที่ใกล้เคียงที่สุดของสถานที่ท่องเที่ยว โรงแรม ร้านอาหาร ที่มีในฐานข้อมูลและ (4) ส่วนของการแสดงผล เป็นการแสดงผลการให้คำแนะนำสถานที่ท่องเที่ยว โรงแรม ร้านอาหาร ให้กับผู้ใช้งานที่ตรงตามความต้องการของผู้ใช้งาน

ระบบถูกสร้างโดยใช้ภาษา PHP เป็นภาษาในการประมวลผลและใช้โปรแกรมระบบจัดการฐานข้อมูลแบบโอเพนซอร์ส MySQL สำหรับจัดเก็บข้อมูลต่างๆนอกจากนี้ได้พัฒนาแอปพลิเคชันให้สามารถใช้งานร่วมกับสมาร์ตโฟน

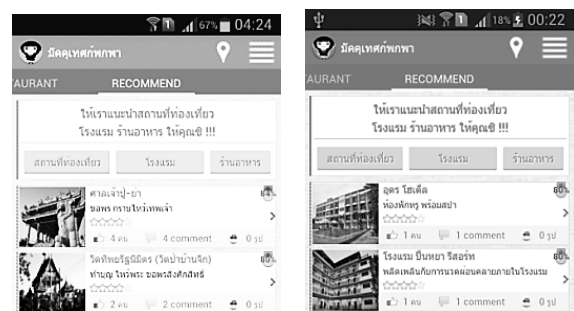
บนระบบปฏิบัติการแอนดรอยด์โดยใช้ฐานข้อมูลร่วมกันทั้งสองระบบ

6. ผลการดำเนินงาน

ผลการดำเนินงานของการพัฒนาโปรแกรมมัลติแพลตฟอร์มบนระบบปฏิบัติการแอนดรอยด์ ในส่วนของระบบแนะนำสถานที่ท่องเที่ยว โรงแรม ร้านอาหาร ผู้ใช้จะต้องกรอกข้อมูลปัจจัยในด้านต่างๆ ของสถานที่ ที่ผู้ใช้งานต้องการพร้อมทั้งให้น้ำหนักความสำคัญของปัจจัยแต่ละข้อ ดังภาพที่ 2 จากนั้นระบบจะแนะนำ สถานที่ท่องเที่ยว โรงแรม ร้านอาหารที่ใกล้เคียงกับที่ผู้ใช้งานต้องการออกมาดังภาพที่ 3



ภาพที่ 2:ขั้นตอนการกรอกข้อมูล



ภาพที่ 3:สถานที่ท่องเที่ยวที่ระบบแนะนำ

ผลการประเมินความพึงพอใจในการใช้งานระบบที่พัฒนาขึ้น แบ่งออกเป็น 2 ส่วน คือ การประเมินความพึงพอใจของระบบโดยผู้เชี่ยวชาญ และการประเมินความพึงพอใจของระบบโดยผู้ใช้งานทั่วไป ทั้ง 5 ด้าน ผลการประเมินความพึงพอใจของผู้เชี่ยวชาญในการใช้ระบบซึ่งค่าเฉลี่ยโดยรวมของทุกด้านเท่ากับ 4.23 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.56 และผล

การประเมินความพึงพอใจของผู้ใช้ทั่วไป แต่ละด้านซึ่งค่าเฉลี่ยโดยรวมของทุกด้านเท่ากับ 4.27 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.60 ดังนั้นระบบที่พัฒนาขึ้น ผู้เชี่ยวชาญและผู้ใช้ทั่วไปมีความเห็นว่าการใช้งานของระบบโดยรวมทุกด้านอยู่ในระดับดี

7. สรุปผลและข้อเสนอแนะ

แอปพลิเคชัน สำหรับแนะนำและให้สารสนเทศเกี่ยวกับสถานที่ท่องเที่ยว โรงแรม และ ร้านอาหารจากตำแหน่งพิกัดภูมิศาสตร์ของสถานที่เหล่านั้น ซึ่งแอปพลิเคชันนี้ทำงานบนระบบปฏิบัติการแอนดรอยด์ ที่ผู้ใช้ที่เคยไปท่องเที่ยวยังสถานที่ต่างๆ สามารถทำการแนะนำสถานที่แห่งใหม่เข้าสู่ระบบ หลังจากนั้นต้องผ่านการตรวจสอบความเหมาะสมจากผู้ดูแลระบบก่อนที่สถานที่แห่งนั้นจะเสนอเป็นสถานที่แนะนำต่อสาธารณชนได้ สำหรับกระบวนการแนะนำแหล่งท่องเที่ยว นั้นจะใช้เทคนิคการหาสมาชิกที่ใกล้ที่สุด (k-Nearest Neighbor) จากข้อมูลเกี่ยวกับข้อคิดเห็นและคะแนนความนิยมด้านความสวยงาม ด้านความคุ้มค่า ด้านการให้บริการ ด้านความสะดวก และด้านเส้นทางการเดินทาง ที่ผู้ท่องเที่ยวคนอื่นๆ ได้ให้คะแนนความนิยมไว้เทียบกับความนิยมของผู้ใช้ที่กำลังใช้งานอยู่ จากการประเมินความพึงพอใจของแอปพลิเคชัน โดยกลุ่มตัวอย่างผู้ใช้งานทั่วไป มีความพึงพอใจอยู่ในระดับดี

สำหรับการศึกษาน่าสนใจในอนาคตคือการพัฒนาคุณสมบัติด้านอื่นประกอบ เช่น การท่องเที่ยวเชิงอนุรักษ์ การท่องเที่ยวเชิงวัฒนธรรม เป็นต้น และอาจจะศึกษาพิจารณาคุณลักษณะด้านอื่นเพิ่มเติม เพื่อให้การแนะนำสถานที่ท่องเที่ยวตอบสนองต่อผู้ใช้ได้มากขึ้น

8. กิตติกรรมประกาศ

ขอขอบคุณสถาบันวิจัยมหาวิทยาลัยราชภัฏอุดรธานี ที่สนับสนุนทุนวิจัย

เอกสารอ้างอิง

- [1] G. Adomavicius and Y. Kwon, "New Recommendation Techniques for Multicriteria Rating Systems," IEEE Intelligence Systems, vol. 22, No. 3, pp.48-55, 2007.
- [2] L. Peterson, "K-nearest neighbor," Scholarpedia, 4(2):1883, 2009. [online] Available from : www.scholarpedia.org/article/-K-nearest_neighbor.
- [3] Marko Balabanovic, Yoav Shoham. (1997). "Fab: Content-based, collaborative recommendation," [ออนไลน์]. ค้นเมื่อ 17 มิถุนายน 2556. จาก : <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.108.4092>
- [4] เกวลิน อัครทรัพย์สาคร และ ณัฐวิ อุดกฤษฎ์, "ระบบช่วยแนะนำในการเลือกซื้อเครื่องประดับผ่านเครือข่ายอินเทอร์เน็ตกรณีศึกษา ร้านเท็นแกลลอรี่", The 5th National Conference on Computing and Information Technology, ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2552.
- [5] เกวลิน อัครทรัพย์สาคร และ ณัฐวิ อุดกฤษฎ์, "ระบบให้คำแนะนำในการเลือกซื้อเลนส์สำหรับกล้องดีเอสแอลอาร์ โดยใช้เทคนิค Content-based Filtering และ Collaborative Filtering", The 7th National Conference on Computing and Information Technology, ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2554.
- [6] วิจิตร ตันทสุทธิ์, วันชัย ริจิรวนิช และศิริจันทร์ ทองประเสริฐ. (2540). การวิจัยดำเนินงาน. กรุงเทพฯ: ซีเอ็ดยูเคชั่น.
- [7] วิภาวรรณ สิงห์พริ้ง. (2541). การวิจัยการดำเนินงาน. กรุงเทพฯ: ภาควิชาวิศวกรรมอุตสาหการ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

การเปรียบเทียบความเร็วในการประมวลผลระหว่างฐานข้อมูลเชิงสัมพันธ์ และฐานข้อมูลไม่สัมพันธ์แบบเอกสาร

Analysis of Data Processing Performance between Relational and Non-Relational Databases of Documents

^{1,2}ประกายมาศ ศรีสุขทักษิณ (Pragaimas Srisuktaksin)¹ และผุสดี บุญรอด (Pudsadee Boonrawd)²

ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
pragaimas@gmail.com¹, pudsadee@kmutnb.ac.th²

บทคัดย่อ

ในปัจจุบันข้อมูลข่าวสารมีปริมาณเพิ่มขึ้นอย่างรวดเร็วจนทำให้เกิดขีดความสามารถของฐานข้อมูลเชิงสัมพันธ์ (Relational Database) การบริหารจัดการข้อมูลที่ดีจึงเป็นสิ่งสำคัญ ดังนั้นฐานข้อมูลไม่สัมพันธ์ (Non Relational Database) จึงเริ่มเข้ามามีบทบาทกับระบบสารสนเทศในปัจจุบันมากยิ่งขึ้น งานวิจัยนี้ได้ทำการทดสอบความเร็วในการประมวลผลของฐานข้อมูลไม่สัมพันธ์แบบเอกสาร ทั้งแบบประมวลผลเครื่องเดียวและประมวลผลแบบกระจายเปรียบเทียบความเร็วกับฐานข้อมูลเชิงสัมพันธ์ด้านการเขียน อ่าน แก้ไขและลบข้อมูล ผลการทดสอบพบว่าหากปริมาณข้อมูลหรือจำนวนเธรด (Thread) มีน้อยความเร็วที่ใช้ในการประมวลผลข้อมูลของฐานข้อมูลแต่ละแบบจะใช้เวลาใกล้เคียงกัน แต่เมื่อปริมาณข้อมูลหรือจำนวนเธรดเพิ่มขึ้นจะยังเห็นความแตกต่างได้อย่างชัดเจน ฐานข้อมูลไม่สัมพันธ์ทั้งสองแบบใช้เวลาน้อยกว่าฐานข้อมูลเชิงสัมพันธ์หลายเท่า โดยเฉพาะการลบข้อมูล ซึ่งพบว่าสามารถลบข้อมูลได้เร็วกว่าฐานข้อมูลเชิงสัมพันธ์ถึงเกือบ 70 เท่า ฐานข้อมูลไม่สัมพันธ์จึงมีประโยชน์มากสำหรับระบบสารสนเทศขนาดใหญ่ซึ่งมีจำนวนการเข้าใช้งานปริมาณมาก และต้องการเวลาในการประมวลผลที่รวดเร็ว

คำสำคัญ: ฐานข้อมูลเชิงสัมพันธ์ ฐานข้อมูลไม่สัมพันธ์แบบเอกสาร ฐานข้อมูลไม่สัมพันธ์ ภาษาสอบถามแบบไม่มีโครงสร้าง ข้อมูลขนาดใหญ่

Abstract

Recently, data and information has been rapidly increasing which is now breaking the maximum limit of relational databases. This, better management of this

tremendous amount of data has become necessary at this time. Therefore, non-relational databases are playing a more important role in the modern work system. This research examines the efficiency of non-relational databases with document database processing and distribution processing. The speed of non-relational database processing was compared with relational database processing in many aspects including writing, reading, updating, and erasing processes. The results indicate that there are no significant differences in the processing time when the data or threads are small. In the other hand, the time cost becomes obviously larger when the data or threads become bigger. Both types of the non-relational databases consume less time and cost than the relational database, especially for erasing process. The erasing process in the non-relational database is 70 times faster than the relational database. Hence, the non-relational database is suitable for a system which has a huge database or high traffic load and it also exhibits swift processing times.

Keywords: Relational Database, Non-Relational with Document Database, Non-Relational Database, NoSQL, Big Data.

1. บทนำ

ข้อมูลข่าวสารในปัจจุบันมีปริมาณเพิ่มขึ้นอย่างรวดเร็ว จนทำให้เกิดข้อมูลปริมาณมาก หรือเรียกว่า “ข้อมูลขนาดใหญ่ (Big Data)” และจากผลการสำรวจข้อมูลของ International Data Corporation (IDC) พบว่าปริมาณข้อมูลตั้งแต่ปีพ.ศ. 2550

ถึงพ.ศ. 2553 มีจำนวนเพิ่มมากขึ้นกว่าเดิมถึง 25 เท่า [1] และมีแนวโน้มเพิ่มขึ้นอย่างต่อเนื่อง ระบบงานที่ถูกพัฒนาด้วยฐานข้อมูลเชิงสัมพันธ์ (Relational Database) ซึ่งมีการจัดเก็บข้อมูลในรูปแบบตารางความสัมพันธ์ในลักษณะแถวและคอลัมน์ทำให้ง่ายต่อการทำความเข้าใจและประยุกต์ใช้งาน [2] แต่การประมวลผลข้อมูลจำนวนมากอาจทำให้ช้าและต้องใช้งบประมาณเป็นจำนวนมากในการเพิ่มประสิทธิภาพของฐานข้อมูล ดังนั้นฐานข้อมูลไม่สัมพันธ์ (Non-Relational Database) จึงกำลังเป็นที่น่าสนใจสำหรับองค์กรทางธุรกิจในปัจจุบันเนื่องจากมีต้นทุนต่ำในการพัฒนา และถูกนำมาใช้จัดการกับข้อมูลปริมาณมาก องค์กรใหญ่ ๆ อย่างเช่น Google, Twitter และ Facebook เป็นต้น มีการนำฐานข้อมูลไม่สัมพันธ์มาประยุกต์ใช้งานกับระบบเพื่อรองรับปริมาณของข้อมูลที่มีแนวโน้มเพิ่มขึ้นอีกในอนาคต

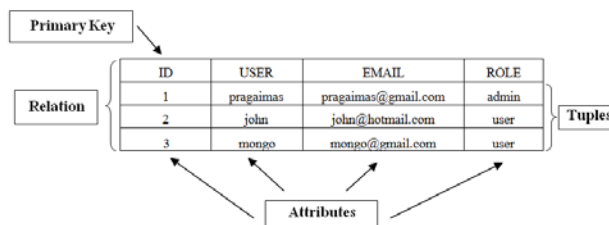
ดังนั้นงานวิจัยนี้มีวัตถุประสงค์เพื่อทดสอบคุณสมบัติเรื่องความเร็วในการประมวลผลของฐานข้อมูล MongoDB ซึ่งเป็นฐานข้อมูลไม่สัมพันธ์แบบเอกสารที่มีลักษณะ โครงสร้างใกล้เคียงกับฐานข้อมูลเชิงสัมพันธ์ โดยแบ่งออกเป็น 2 แบบคือแบบประมวลผลเครื่องเดียวและประมวลผลแบบกระจาย (Sharding) เปรียบเทียบความเร็วในการอ่าน (Read) เขียน (Write) แก้ไข (Update) และการลบ (Delete) กับฐานข้อมูล MySQL ซึ่งเป็นฐานข้อมูลเชิงสัมพันธ์ โดยมีปัจจัยในการทดสอบคือจำนวนข้อมูลและจำนวนการติดต่อฐานข้อมูลที่แตกต่างกัน โดยบทความนี้ประกอบด้วยหัวข้อที่ 1 บทนำ หัวข้อที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง หัวข้อที่ 3 วิธีการดำเนินงาน หัวข้อที่ 4 แสดงผลการดำเนินงานและหัวข้อที่ 5 สรุปผลการทดสอบรวมถึงข้อเสนอแนะ

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ฐานข้อมูลเชิงสัมพันธ์ (Relational Database)

ฐานข้อมูลเชิงสัมพันธ์ เป็นฐานข้อมูลที่ใช้โมเดลเชิงสัมพันธ์เป็นแบบจำลองที่มีการใช้ทั่วไปในรูปแบบของตาราง (Table) ในแต่ละตารางแบ่งออกเป็นแถว (Record) ในแต่ละแถวแบ่งเป็นคอลัมน์ (Column) ในทางทฤษฎีจะมีศัพท์เฉพาะแตกต่างกันไป ซึ่งเกิดจากทฤษฎีทางคณิตศาสตร์เรื่องเซต (Set) ซึ่งง่ายต่อการทำความเข้าใจและประยุกต์ใช้งาน จึงได้รับความ

นิยามที่สุด[2] ซึ่งมีส่วนประกอบ โครงสร้างพื้นฐานของฐานข้อมูลเชิงสัมพันธ์ประกอบด้วย รีเลชัน (Relation) แอททริบิวต์ (Attribute) ทูเปิล (Tuple) หรือแถว (Record) และคีย์หลัก (Primary Key) [3] ดังภาพที่ 1



ภาพที่ 1: โครงสร้างฐานข้อมูลเชิงสัมพันธ์

2.2 ฐานข้อมูลไม่สัมพันธ์ (Non-Relational Database)

มีแนวความคิดตรงกันข้ามกับฐานข้อมูลเชิงสัมพันธ์ จะไม่มีความสัมพันธ์ระหว่างข้อมูลและใช้ภาษาสอบถามแบบไม่มีโครงสร้าง (NoSQL) ในการจัดการกับข้อมูล และโครงสร้างถูกออกแบบมาเพื่อรองรับข้อมูลที่มีปริมาณมาก มี 4 ประเภทตามลักษณะการจัดเก็บข้อมูลที่แตกต่างกันนี้ [1]

2.2.1 ฐานข้อมูลแบบใช้คีย์อ้างอิง (Key Value Store) เก็บข้อมูลในลักษณะของคีย์และผลลัพธ์ โดยอ้างอิงข้อมูลผ่านทางค่าที่เป็นลักษณะเฉพาะไม่ซ้ำกับข้อมูลตัวอื่น ๆ (Key) เพื่อดึงถึงผลลัพธ์ออกมา และลักษณะ โครงสร้างของฐานข้อมูลมีลักษณะอย่างง่าย และรองรับข้อมูลปริมาณมากรวมถึงการแก้ปัญหา การเกิดภาวะพร้อมกัน (Concurrency) เช่น BerkeleyDB และ Tokyo Tyrant เป็นต้น

2.2.2 ฐานข้อมูลแบบตารางขนาดใหญ่ (Big Table) เก็บข้อมูลกระจายตามเซิร์ฟเวอร์ เป็นตารางขนาดใหญ่ที่ Google พัฒนาขึ้นมาใช้กับระบบมี 3 มิติ คือนอกจากแถวกับคอลัมน์แล้ว ยังมีมิติของเวลา เพื่อแสดงข้อมูลเฉพาะบางช่วงเวลา

2.2.3 ฐานข้อมูลแบบเอกสาร (Document Databases) เก็บข้อมูลในรูปแบบของเอกสาร มีโครงสร้างทั้งลึกและกว้างในแต่ละแถวเหมือนกับอาเรย์หลายมิติไม่แบนราบเหมือน MySQL 1 แถว ซึ่งมีแค่มิติเดียวเท่านั้น เช่น MongoDB และ Couch DB เป็นต้น [4]

2.2.4 ฐานข้อมูลแบบกราฟ (Graph Databases) ข้อมูลถูกจัดเก็บในลักษณะของความสัมพันธ์ ซึ่งจะมีโหนดเป็นตัวแทนของข้อมูลเชื่อมความสัมพันธ์กัน เช่น Graph Database

2.3 งานวิจัยที่เกี่ยวข้อง

งานวิจัยมีการนำฐานข้อมูลเชิงสัมพันธ์เปรียบเทียบประสิทธิภาพกับฐานข้อมูลไม่สัมพันธ์เช่น ทดสอบความเร็วในการประมวลผลของ MySQL กับ MongoDB [5] หรือการทดสอบการทำงานของ Oracle NoSQL [6] เปรียบเทียบการทำงาน คุณความยืดหยุ่นและความถูกต้องโดยไม่คำนึงถึงคุณสมบัติของฐานข้อมูลเชิงสัมพันธ์ ACID วัตถุประสงค์ของการอ่าน การเขียน และการแก้ไข ผลปรากฏว่าฐานข้อมูลไม่สัมพันธ์สามารถประมวลผลข้อมูลได้เร็วกว่าฐานข้อมูลเชิงสัมพันธ์หลายเท่า และเปรียบเทียบคุณสมบัติของฐานข้อมูลทั้ง 2 แบบ [7] และสรุปผลในแต่ละด้าน เช่น ความเร็วในการประมวลผลข้อมูล ความยืดหยุ่น การทำข้อมูลประมวลรายการ (Transaction) หรือลักษณะการใช้ดัชนี (Index) เป็นต้น [8], [9] หรือมีการนำ MySQL ใช้งานผสมผสานกับ Cassandra โดยการนำข้อมูลเก็บไว้ใน MySQL และ Lucene ซึ่งมีการเพิ่มความสามารถในการเก็บข้อมูลเพิ่มเติมอีก จึงทำให้งานที่ได้มีคุณสมบัติทั้งด้าน ACID และด้านความเร็วในการประมวลผลจากฐานข้อมูลไม่สัมพันธ์ [10]

การนำเสนอสถาปัตยกรรมการออกแบบฐานข้อมูลแบบไม่สัมพันธ์โดยเน้นที่ประสิทธิภาพการประมวลผลฐานข้อมูลแบบกระจาย เปรียบเทียบความเร็วของจำนวนเครื่องเซิร์ฟเวอร์ ปรากฏว่ายิ่งจำนวนเซิร์ฟเวอร์เพิ่มขึ้นก็ยิ่งทำให้ใช้เวลาน้อยลง [11] หรือการใช้ฐานข้อมูล Cassandra และ Neo4j สร้างโมเดลต้นแบบในการพัฒนาระบบ ทดลองสุ่มตัวอย่างข้อมูลนำเข้า และจัดกลุ่มรูปแบบคำสั่งที่มีความคล้ายคลึงกัน หรือเปรียบเทียบคุณสมบัติฐานข้อมูลไม่สัมพันธ์ Google BigTable, Amazon SimpleDB, Apache CouchDB, MongoDB, Cassandra และ Hbase [12] หรือทดสอบปรับขยายความสามารถของฐานข้อมูลไม่สัมพันธ์ และดูผลการประมวลผลของหน่วยประมวลผลหลัก ค่าเฉลี่ยของเวลา ค่าหน่วยความจำที่ใช้ โดยในส่วนของผู้ให้บริการจะใช้เทคนิคการทำคลัสเตอร์ (Cluster) ช่วยในการประมวลผลของฐานข้อมูล ผลที่ได้ปรากฏว่าฐานข้อมูลไม่สัมพันธ์สามารถประมวลผลข้อมูลขนาดใหญ่ได้อย่างรวดเร็ว ไม่เปลืองซีพียูและเวลาที่ใช้งาน [13]

3. วิธีการดำเนินการวิจัย

เครื่องสำหรับทดสอบในการดำเนินงานวิจัยใช้หน่วยประมวลผลกลาง (CPU) แบบ Intel Core i5 2.27 GHz หรือสูงกว่า หน่วยความจำหลัก (RAM) มีความจุไม่น้อยกว่า 8 GB ฮาร์ดดิสก์มีความจุไม่น้อยกว่า 500 GB และแบ่งการดำเนินงานเป็น 4 หัวข้อดังนี้

3.1 ฐานข้อมูลสำหรับทดสอบ

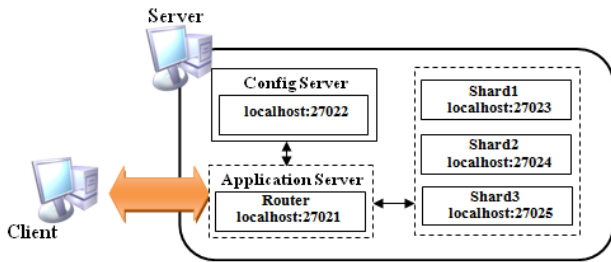
3.1.1 ฐานข้อมูลเชิงสัมพันธ์ใช้ MySQL เวอร์ชัน 5.0.51b ซึ่งมีโครงสร้างฐานข้อมูลดังตารางที่ 1

ตารางที่ 1: โครงสร้างฐานข้อมูลเชิงสัมพันธ์ที่ใช้ทดสอบ

ชื่อข้อมูล	ชนิดข้อมูล	ขนาด	หมายเหตุ
id	Varchar	10	Primary Key
col1	Varchar	10	-
col2	Varchar	10	-
col3	Varchar	10	-
col4	Varchar	10	-

3.1.2 ฐานข้อมูลไม่สัมพันธ์ใช้ MongoDB เวอร์ชัน 2.4.5 แบบประมวลผลเครื่องเดียว ซึ่งมีโครงสร้างเอกสารดังภาพที่ 2

3.1.3 ฐานข้อมูลไม่สัมพันธ์ใช้ MongoDB เวอร์ชัน 2.4.5 และแบ่งการประมวลผลออกไปยัง 3 เครื่องย่อย (Sharding) คือ Shard1 (localhost:27023), Shard2 (localhost:27024) และ Shard3 (localhost:27025) โดยไคลเอนต์ (Client) จะส่งประมวลผลโปรแกรม ไปยังส่วนของแอปพลิเคชันเซิร์ฟเวอร์ (Application Server) จะทำหน้าที่เปรียบเสมือนเป็นเราเตอร์ โดยจะอ่านค่าคอนฟิก (Config) จากคอนฟิกเซิร์ฟเวอร์ (Config Server) และทำการแจกจ่ายงานไปยัง 3 เครื่องย่อยซึ่งเป็นการกระจายของการจัดเก็บข้อมูลในหลายเครื่อง และเป็นการขยายฐานข้อมูลแบบแนวนอน (Horizontal Scaling) เพื่อรองรับการขยายระบบในอนาคต เนื่องจากคอมพิวเตอร์เครื่องเดียวอาจจะไม่เพียงพอในการอ่าน เขียน และจัดเก็บข้อมูลขนาดใหญ่ ดังภาพที่ 2 [14] และมีโครงสร้างเอกสารซึ่งมีลักษณะคล้ายกับฐานข้อมูลเชิงสัมพันธ์ ดังภาพที่ 3



ภาพที่ 2: สถาปัตยกรรม Sharding MongoDB ที่ใช้ทดสอบ

```

Json Text of Resulting Documents:
/* 0 */
{
  "id" : 1
  "col1" : "xjdhojkfdl"
  "col2" : "dkjufcepww"
  "col3" : "klopmsjhjev"
  "col4" : "wmlkmaqpxz"
}
    
```

ภาพที่ 3: ตัวอย่างโครงสร้างเอกสาร MongoDB ที่ใช้ทดสอบ

3.2 การเพิ่มประสิทธิภาพให้กับฐานข้อมูล

3.2.1 ฐานข้อมูล MySQL คณะผู้วิจัยได้เพิ่มประสิทธิภาพ โดยการสร้างดัชนีตามคำสั่งดังนี้

```
CREATE INDEX indx_id ON test (id);
```

3.2.2 ฐานข้อมูล MongoDB ทั้งแบบประมวลผลเครื่องเดียวและประมวลผลแบบกระจาย คณะผู้วิจัยได้เพิ่มประสิทธิภาพ โดยการสร้างดัชนีตามคำสั่งดังนี้

```
db.test.ensureIndex({ id: 1 })
```

3.3 การออกแบบการทดสอบฐานข้อมูล ถูกพัฒนาด้วยภาษาจาวาและแบ่งการทดสอบออกเป็น 4 ด้านคือ

3.3.1 ทดสอบความสามารถในการอ่าน (Read) ของฐานข้อมูลแต่ละแบบ โดยการอ่านข้อมูลจำนวน 500, 5,000, 50,000 และ 500,000 แถว จากฐานข้อมูลที่มีจำนวนเท่ากับ 10,000,000 แถว ซึ่งจะมีจำนวนการติดต่อฐานข้อมูลแต่ละจำนวนที่แตกต่างกันดังนี้ 1 เกรด 10 เกรด และ 100 เกรด

3.3.2 ทดสอบความสามารถในการเขียน (Insert) ของฐานข้อมูลแต่ละแบบ โดยการเขียนข้อมูลจำนวน 1,000, 10,000, 100,000, 1,000,000 และ 10,000,000 แถว ซึ่งจะมีจำนวนการติดต่อฐานข้อมูลแต่ละจำนวนที่แตกต่างกันดังนี้ 1 เกรด 10 เกรด และ 100 เกรด

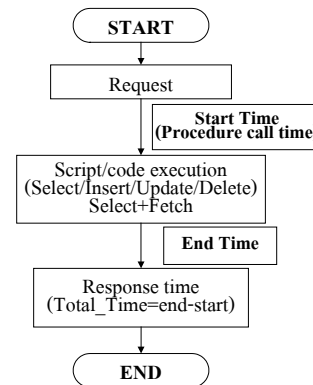
3.3.3 ทดสอบความสามารถในการแก้ไข (Update) ของฐานข้อมูลแต่ละแบบ โดยการแก้ไขข้อมูลจำนวน 500, 5,000,

50,000 และ 500,000 แถว จากฐานข้อมูลที่มีจำนวนเท่ากับ 10,000,000 แถว ซึ่งจะมีจำนวนการติดต่อฐานข้อมูลแต่ละจำนวนที่แตกต่างกันดังนี้ 1 เกรด 10 เกรด และ 100 เกรด

3.3.4 ทดสอบความสามารถในการลบ (Delete) ของฐานข้อมูลแต่ละแบบ โดยการลบข้อมูลจำนวน 500, 5,000, 50,000 และ 500,000 แถว จากฐานข้อมูลที่มีจำนวนเท่ากับ 10,000,000 แถว ซึ่งจะมีจำนวนการติดต่อฐานข้อมูลแต่ละจำนวนที่แตกต่างกันดังนี้ 1 เกรด 10 เกรด และ 100 เกรด

3.4 การวัดค่าเวลาเพื่อทดสอบความเร็วในการประมวลผล

กระบวนการวัดค่าเวลาการประมวลผลของฐานข้อมูลแต่ละแบบเริ่มจากการส่งคำร้องขอไปยังฐานข้อมูล และจากนั้นประมวลผลตัวดำเนินการต่าง ๆ ที่มีกรออกแบบการทดลองไว้คือการอ่าน การเขียน การแก้ไข และการลบ บันทึกเวลาหลังการประมวลผลเสร็จจะได้เวลาที่ฐานข้อมูลใช้สำหรับประมวลผลดังภาพที่ 4



ภาพที่ 4: การวัดความเร็วในการประมวลผลข้อมูล

การคำนวณเวลาที่ไว้วัดความเร็วในการประมวลผลของฐานข้อมูลแต่ละแบบมีการคำนวณดังสมการที่ 1

$$Total_Time = End_Time - Begin_Time \quad (1)$$

Total_Time คือ เวลาทั้งหมดที่ใช้ในการประมวลผล

End_Time คือ เวลาที่ใช้ในการ Select และ Fetch ข้อมูล

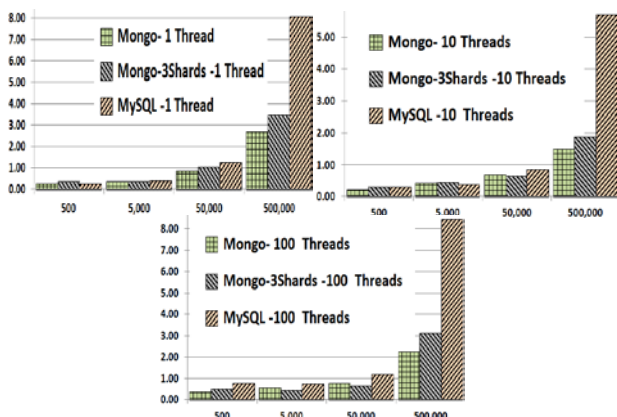
Begin_Time คือ เวลาเริ่มต้นในการประมวลผลข้อมูล

การทดสอบแต่ละกรณีนั้นทำการทดลอง 3 ครั้ง และหาค่าเฉลี่ย

4. ผลการดำเนินงาน

4.1 ผลการทดสอบความสามารถในการอ่านข้อมูล

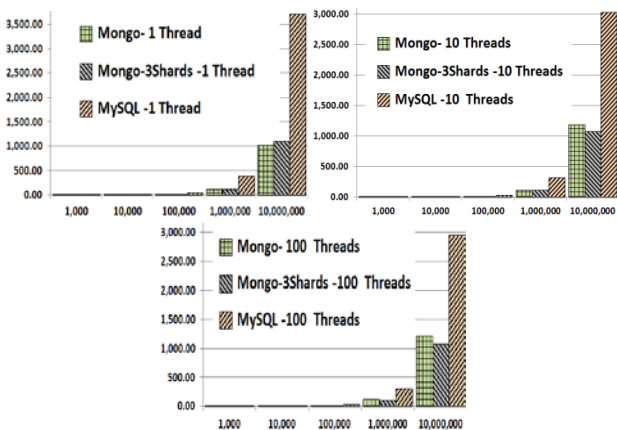
ผลการทดสอบฐานข้อมูลแต่ละแบบพบว่าหากข้อมูลและจำนวนเรดน้อยความสามารถในการอ่านของฐานข้อมูลแต่ละแบบจะไม่แตกต่างกันมาก แต่เมื่อจำนวนข้อมูลหรือจำนวนเรดมีปริมาณเพิ่มขึ้นพบว่า MongoDB แบบประมวลผลเครื่องเดียวและประมวลผลแบบกระจายซึ่งมีความสามารถใกล้เคียงกันทำงานได้เร็วกว่า MySQL ประมาณ 3 เท่า ดังภาพที่ 5



ภาพที่ 5: ผลการทดสอบการอ่านข้อมูล

4.2 ผลการทดสอบความสามารถในการเขียนข้อมูล

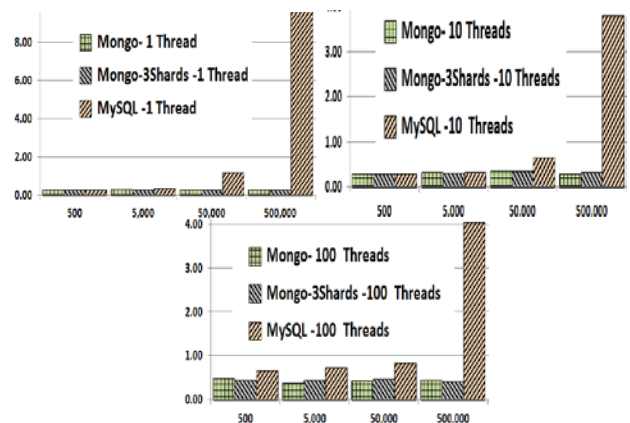
ผลการทดสอบฐานข้อมูลแต่ละแบบพบว่าหากข้อมูลและจำนวนเรดน้อยความสามารถในการเขียนของฐานข้อมูลจะไม่แตกต่างกันมาก แต่เมื่อจำนวนข้อมูลหรือจำนวนเรดเพิ่มมากขึ้นพบว่า MongoDB แบบประมวลผลเครื่องเดียวและประมวลผลแบบกระจายซึ่งมีความสามารถใกล้เคียงกันทำงานได้เร็วกว่า MySQL ประมาณ 3 เท่า ดังภาพที่ 6



ภาพที่ 6: ผลการทดสอบการเขียนข้อมูล

4.3 ผลการทดสอบความสามารถในการแก้ไขข้อมูล

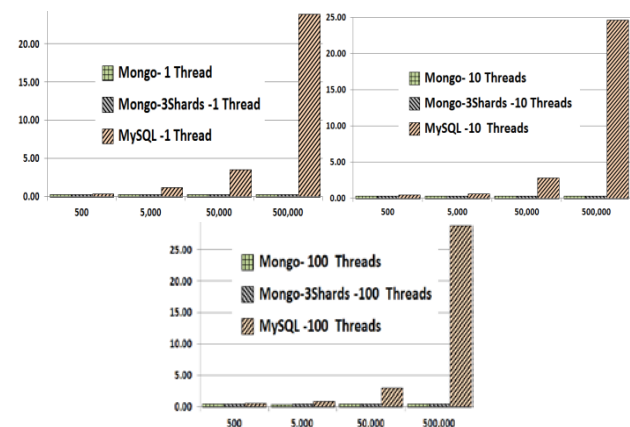
ผลการทดสอบฐานข้อมูลแต่ละแบบพบว่า หากข้อมูลและจำนวนเรดน้อยความสามารถในการแก้ไขของฐานข้อมูลจะไม่แตกต่างกันมาก แต่เมื่อจำนวนข้อมูลหรือจำนวนเรดมีปริมาณเพิ่มขึ้นพบว่า MongoDB แบบประมวลผลเครื่องเดียวและแบบกระจายซึ่งมีความสามารถใกล้เคียงกันจะทำงานได้เร็วกว่า MySQL หลายเท่า จากการทดสอบพบว่าที่จำนวนข้อมูลเท่ากับ 500,000 แถวและจำนวนเรดเท่ากับ 1 เรด MongoDB ทั้งสองแบบทำงานเร็วกว่า MySQL เกือบ 40 เท่า ดังภาพที่ 7



ภาพที่ 7: ผลการทดสอบการแก้ไขข้อมูล

4.4 ผลการทดสอบความสามารถในการลบข้อมูล

ผลการทดสอบฐานข้อมูลแต่ละแบบพบว่า เมื่อเพิ่มจำนวนข้อมูลและจำนวนเรด MongoDB แบบประมวลผลเครื่องเดียวและแบบกระจายใช้เวลาใกล้เคียงกัน ส่วน MySQL เมื่อจำนวนข้อมูลหรือจำนวนเรดมีปริมาณเพิ่มขึ้นก็จะใช้เวลาประมวลผลเพิ่มขึ้นด้วย จากการทดสอบพบว่าที่จำนวนข้อมูลเท่ากับ 500,000 แถวและจำนวนเรดเท่ากับ 100 เรด MongoDB ทั้งสองแบบทำงานเร็วกว่า MySQL เกือบ 70 เท่า ดังภาพที่ 8



ภาพที่ 8: ผลการทดสอบการลบข้อมูล

5. สรุปผลการวิจัยและข้อเสนอแนะ

งานวิจัยนี้ได้ทดสอบเกี่ยวกับความเร็วในการประมวลผลของ MySQL ซึ่งเป็นฐานข้อมูลเชิงสัมพันธ์ MongoDB แบบประมวลผลเครื่องเดียวและ MongoDB แบบประมวลผลกระจาย ซึ่งเป็นฐานข้อมูลไม่สัมพันธ์ จากการทดสอบพบว่าหากข้อมูลหรือจำนวนเรคคอร์ดน้อย ฐานข้อมูลจะมีความสามารถในการประมวลผลใกล้เคียงกัน แต่เมื่อข้อมูลหรือจำนวนเรคคอร์ดมีปริมาณมากขึ้นก็จะเริ่มเห็นความแตกต่างอย่างชัดเจน โดยเฉพาะเรื่องการลบข้อมูลพบว่าปริมาณข้อมูลและจำนวนเรคคอร์ดมีผลน้อยมากกับความเร็วในการลบข้อมูลของ MongoDB ซึ่งประมวลผลเร็วกว่า MySQL หลายเท่า การทำ Sharding MongoDB เพื่อลดภาระการทำงานของเครื่องเซิร์ฟเวอร์ และรองรับการขยายระบบในอนาคต พบว่ามีกระบวนการติดตั้งที่ไม่ซับซ้อน โดยประสิทธิภาพในการประมวลผลใกล้เคียงกับการประมวลผลแบบเครื่องเดียว และมีแนวโน้มใช้เวลาน้อยกว่าแบบประมวลผลเครื่องเดียว หากข้อมูลและจำนวนเรคคอร์ดมีปริมาณมาก ดังนั้นฐานข้อมูลไม่สัมพันธ์จึงมีประโยชน์มากสำหรับระบบที่มีข้อมูลขนาดใหญ่ ต้องการเวลาการประมวลผลที่รวดเร็ว และระบบที่มีการเติบโตของข้อมูลสูง อีกทั้งยังช่วยประหยัดงบประมาณในการปรับเปลี่ยนประสิทธิภาพของฐานข้อมูล งานวิจัยในอนาคตอาจมีการปรับเปลี่ยนประสิทธิภาพของฐานข้อมูลเชิงสัมพันธ์เช่นการปรับเปลี่ยนให้ฐานข้อมูลสามารถรองรับการเข้าถึงฐานข้อมูลเวลาเดียวกันอย่างมีประสิทธิภาพ วิจัยเรื่องการสำรองข้อมูล หรือคุณสมบัติต่าง ๆ ของฐานข้อมูลไม่สัมพันธ์ที่ออกแบบมาเพื่อรองรับข้อมูลขนาดใหญ่ โดยเฉพาะ เพื่อเป็นประโยชน์ต่อผู้ที่สนใจได้นำไปพิจารณาความเหมาะสมกับระบบที่จะพัฒนาขึ้น

เอกสารอ้างอิง

- [1] Clarence J M Tauro and et al., "Comparative Study of the New Generation, Agile, Scalable, High Performance NOSQL Databases", *International Journal of Computer Applications (0975 – 888)*, vol. 48, no 20, pp. 20-24, 2012.
- [2] บุรินทร์ รุจจนพันธุ์ และคณะ, "พัฒนาระบบฐานข้อมูลการประเมินตนเองของหน่วยงาน กรณีศึกษา มหาวิทยาลัยขอนแก่น จังหวัดลำปาง" การประชุมทางวิชาการระดับชาติ ด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ปี 2009 (Proceeding), ครั้งที่ 5, พฤษภาคม 2552.
- [3] สุจินดา ล้อมเจริญผล, "โปรแกรมแปลงแบบจำลองเชิงวัตถุไปสู่ภาษากำหนดโครงสร้างของฐานข้อมูลเชิงสัมพันธ์" ปัญหาพิเศษสาขาวิชาเทคโนโลยีสารสนเทศ ภาควิชาเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2549.
- [4] BeYourCyber Team "MongoDB คืออะไร" Available online at http://meewebfree.com/print_article/general-web-technic/378-what-is-mongodb-database, 2012.
- [5] Parker and et al., "Comparing NoSQL MongoDB to an SQL DB," *Faculty of Computer Science, University of Alabama*, 2013.
- [6] Ashok Joshi and et al., "Oracle NoSQL Database – Scalable, Transactional Key-value Store," *The Second International Conference on Advances in Information Mining and Management*, vol. 11, pp. 15-30, 2012.
- [7] Rabi Prasad Padhy and et al., "RDBMS to NoSQL: Reviewing Some Next-Generation Non-Relational Database's," *Faculty of Computer Science, Berhampur University, India*, 2011.
- [8] Nishtha Jatana and et al., "A Survey and Comparison of Relational and Non-Relational Database," *International Journal of Engineering Research & Technology*, vol. 6, Aug 2012.
- [9] Jiri Schindler, "I/O Characteristics of NoSQL Databases," *NetApp Inc.*, vol.5, pp. 2020-2021, 2012.
- [10] Aaron Schram and Kenneth M. Anderson, "MySQL to NoSQL Data Modeling Challenges in Supporting Scalability," *Faculty of Computer Science, Colorado Boulder University, Colorado, USA*, 2012.
- [11] Fernando G. Tinetti and et al., "Distributed Search on Large NoSQL Databases," *Faculty of Information, UNLP, La Plata, Argentina*, 2011.
- [12] Petter Nasholm, "Extracting Data from NoSQL Databases: A Step towards Interactive Visual Analysis of NoSQL Data," *Master of Science Thesis, Faculty of Computer Science and Engineering, Gothenburg University*, 2012.
- [13] Véronique Benzaken and et al., "Static and Dynamic Semantics of NoSQL Languages," *Paris-Sud University, Orsay, France*, 2013.
- [14] MongoDB, Inc., "MongoDB Documentation Release 2.4.9.," Jan 14, 2014.

การพัฒนาโปรแกรมดูแลสุขภาพสุนัขบนโทรศัพท์มือถือ Development of Dog's Health Care Application on Mobile Phone

ณัฐวดี หงษ์บุญมี(Nattavadee Hongboonmee)¹ วรินยา สงแจ้ง(Wirinya Songcheang)¹ และอุษา ชูทอง(Usa Choothong)¹

¹ภาควิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยนเรศวร

nattavadeeho@nu.ac.th

บทคัดย่อ

การศึกษานี้ เป็นงานวิจัยประยุกต์ มีจุดประสงค์เพื่อศึกษาและพัฒนาแอปพลิเคชันเพื่อช่วยในการดูแลสุขภาพสุนัขบนโทรศัพท์มือถือระบบปฏิบัติการแอนดรอยด์ แอปพลิเคชันนี้พัฒนาด้วยภาษา Java และใช้โปรแกรม SQLite เป็นฐานข้อมูล โดยเน้นที่การเก็บข้อมูลประวัติสุนัข เพื่อแจ้งเตือนการฉีดวัคซีนให้ตรงตามเวลาที่กำหนด วิธีการดูแลสุขภาพในแต่ละสายพันธุ์ อาหารที่สุนัขไม่ควรรับประทาน การวินิจฉัยโรคเบื้องต้นจากการสังเกตอาการสุนัข ผลการพัฒนาแอปพลิเคชันนี้จะช่วยอำนวยความสะดวกต่อผู้เลี้ยงสุนัข และช่วยให้การดูแลสุขภาพแต่ละสายพันธุ์ได้ถูกต้อง และมีประสิทธิภาพมากขึ้น จากการนำแอปพลิเคชันไปประเมินกับผู้ใช้ที่เกี่ยวข้อง พบว่าความพึงพอใจของผู้ใช้ในภาพรวมได้ค่า 4.35 จากผู้ใช้ 33 คน และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.60 ซึ่งถือว่าแอปพลิเคชันนี้มีความพึงพอใจ สามารถนำไปใช้ได้อยู่ในระดับดี

คำสำคัญ: การดูแลสุขภาพสุนัข แอนดรอยด์ โปรแกรมบนมือถือ

Abstract

This research is applied research, aimed to study and development of dog's health care application on mobile phone. This application develop with Java language and database management by SQLite, with emphasis on data collection of dogs, to alert the vaccination at the time. And how to take care of dogs in each species. Feed dogs should not eat. An

initial diagnosis of observation in dogs. The development of this application will help facilitate the dog. And to help take care of each dog breed is correct and effective. The applications were evaluated with users involved. The satisfaction of users found that overall is 4.35 from users 33 people, and the standard deviation is 0.60, which considered this application are satisfied can be used at a good level.

Keywords: Dog's health care, Android, Mobile application.

1. บทนำ

ในปัจจุบัน สุนัขถือเป็นสัตว์เลี้ยงที่มนุษย์นิยมเลี้ยงกันมากที่สุด ทั้งที่เลี้ยงไว้เฝ้าบ้าน เลี้ยงไว้เป็นเพื่อนแก้เหงา ฯลฯ การดูแลสุขภาพสุนัขให้ถูกวิธีเพื่อให้สุนัขมีสุขภาพที่แข็งแรง ร่าเริง จึงเป็นสิ่งสำคัญสำหรับคนที่รักสุนัขเป็นอย่างยิ่ง เพราะสุนัขก็เหมือนกับมนุษย์ ต้องการความรักและการดูแลเอาใจใส่ อีกทั้งการได้รับวัคซีนเป็นสิ่งจำเป็นของสุนัข เพราะหากสุนัขไม่ได้รับวัคซีนตามช่วงอายุที่กำหนดไว้ นั่นอาจส่งผลให้สุนัขมีภูมิคุ้มกันโรคต่างๆ น้อยลง สุขภาพอ่อนแอไม่แข็งแรง ป่วยและติดเชื้อง่ายกว่าสุนัขที่ได้รับวัคซีน แต่เนื่องจากสังคมในปัจจุบัน มนุษย์ใช้เวลาในการทำงานมากขึ้น ทำให้เจ้าของสุนัขอาจลืมว่าถึงกำหนดการได้รับวัคซีนของสุนัขตัวนั้นๆ หรือบางครั้งเจ้าของอาจมีสุนัขหลายตัว ทำให้ไม่สามารถจดจำกำหนดวันฉีดวัคซีนของสุนัขได้ครบถ้วน หรือในบางกรณีเจ้าของ อาจทำสมุดประจำตัวของสุนัขหาย จึงทำให้ไม่ทราบวันนัดฉีดวัคซีน

ของสุนัข นอกจากนั้น ผู้ที่เลี้ยงสุนัขบางคนอาจจะไม่ทราบวิธีการดูแลสุนัขที่ถูกต้อง เนื่องจากลักษณะนิสัยของสุนัขแต่ละพันธุ์ไม่เหมือนกัน จึงมีวิธีการดูแลที่แตกต่างกันออกไป และอีกหนึ่งเรื่องที่สำคัญของการดูแลเลี้ยงดูสุนัขนั้น คือ เรื่องโรคของสุนัข เพราะมนุษย์ไม่สามารถเข้าใจภาษาของสุนัขได้ ดังนั้นเมื่อสุนัขมีอาการผิดปกติ ซึ่งอาจเป็นอาการเริ่มแรกของโรคบางอย่าง เจ้าของจึงต้องใส่ใจและหมั่นสังเกตความผิดปกติของสุนัข เพื่อจะได้พาไปพบสัตวแพทย์ได้ทันเวลา อาหารของสุนัขก็เป็นอีกหนึ่งเรื่องที่สำคัญเช่นกัน เนื่องจากคนส่วนใหญ่ มักคิดว่าสุนัขกินอาหารได้เหมือนกับมนุษย์ จึงคิดว่าสิ่งที่มนุษย์รับประทานได้ สุนัขก็สามารถกินได้เช่นกัน แต่นั่นเป็นความคิดที่ผิด เพราะอาหารของมนุษย์บางอย่างเป็นพิษต่อสุนัข หากสุนัขกินเข้าไปอาจทำให้สุนัขมีอาการเวียนหัว อาเจียน เจ็บป่วยหรืออาจเสียชีวิตได้

ด้วยความสำคัญของปัญหาดังกล่าว ทำให้คณะผู้ศึกษาวิจัยมีแนวคิดที่จะศึกษา ออกแบบและพัฒนาแอปพลิเคชันเพื่อช่วยในการดูแลสุขภาพสุนัข โดยมีการเก็บข้อมูลประวัติสุนัข เพื่อใช้คำนวณช่วงอายุที่จะต้องพาสุนัขไปรับบริการฉีดวัคซีน แล้วทำการแจ้งเตือนเป็นแมสเสจขึ้นในโทรศัพท์ และยังมีข้อมูลการดูแลสุนัข (กรณีศึกษา 5 สายพันธุ์ได้แก่พันธุ์ชิสุ ปอมเมอเรเนียน พุดเดิ้ล ชิววา และพันธุ์บางแก้ว) อย่างถูกต้องวิธี อาหารที่สุนัขไม่ควรรับประทาน และการวินิจฉัยโรคเบื้องต้นจากการสังเกตอาการสุนัข

จากความก้าวหน้าของเทคโนโลยี SmartPhone ซึ่งเป็นโทรศัพท์มือถือ ที่สามารถทำงานได้มากกว่าโทรศัพท์มือถือทั่วไป โดยเฉพาะ Smartphone บนระบบปฏิบัติการแอนดรอยด์ และจากปัญหาที่เกิดจากการดูแลสุขภาพสุนัข ทำให้คณะผู้ศึกษา มีความสนใจที่จะพัฒนาแอปพลิเคชันการดูแลสุขภาพสุนัขบนโทรศัพท์ระบบปฏิบัติการแอนดรอยด์ เพื่ออำนวยความสะดวกให้กับเจ้าของสุนัข และช่วยให้การดูแลสุนัขเป็นไปอย่างถูกต้อง มีประสิทธิภาพมากขึ้น

วัตถุประสงค์ของการศึกษา

1. เพื่อศึกษาการดูแลสุขภาพของสุนัข

2. เพื่อพัฒนาแอปพลิเคชันการดูแลสุขภาพของสุนัขบนโทรศัพท์มือถือ
3. เพื่อประเมินแอปพลิเคชันการดูแลสุขภาพของสุนัขบนโทรศัพท์มือถือ

2. เอกสารและงานวิจัยที่เกี่ยวข้อง

2.1 การพัฒนาแอปพลิเคชันบนระบบปฏิบัติการแอนดรอยด์

แอนดรอยด์ [1] เป็นระบบปฏิบัติการที่ทำงานบนอุปกรณ์พกพา Mobile Device ซึ่งทำงานได้ทั้ง Smartphone หรือ Tablet โดยมี Google เป็นเจ้าของหลัก แต่ทั้งนี้ความสำเร็จของแอนดรอยด์ ไม่ใช่เพียงแค่ Google แต่ยังมีพันธมิตรทางธุรกิจ ด้านการสื่อสารโทรคมนาคม รวมไปถึงบริษัทผลิตแอปพลิเคชันและบริษัทผู้ผลิตอุปกรณ์ต่างๆรวมตัวกันก่อตั้งองค์กร ที่มีชื่อว่า Open handset alliance โดยระบบแอนดรอยด์นั้น อยู่ในกลุ่มของ Open Source Project นั่นคือ มีการเปิดเผย Sourcecode ให้ นักพัฒนาทั่วโลกสามารถพัฒนาต่อยอดได้ นอกจากนั้นระบบปฏิบัติการแอนดรอยด์ ยังมีชุมชนนักพัฒนาซอฟต์แวร์ขนาดใหญ่ ที่คอยสร้างสรรค์แอปพลิเคชันต่างๆ ให้กับโทรศัพท์มือถือที่มีระบบปฏิบัติการแอนดรอยด์ โดยปัจจุบันมีโปรแกรมต่างๆ มากมาย และในการเขียน หรือการพัฒนาโปรแกรมบนแอนดรอยด์ สามารถเขียนได้โดยใช้ภาษาจาวา เพื่อควบคุมอุปกรณ์ต่างๆโดยผ่านทาง Google-developed Java libraries

2.2 SQLite

SQLite เป็นฐานข้อมูลขนาดเล็ก [2] เช่นเดียวกับฐานข้อมูล Microsoft Access สิ่งที่ทำให้ SQLite แตกต่างกับ Microsoft Access คือ ไม่มีค่าลิขสิทธิ์ การติดตั้งง่าย และไม่จำกัดระบบปฏิบัติการทั้งวินโดวส์ แมคอินทอช และลินุกซ์ สำหรับฐานข้อมูลของ SQLite เป็นลักษณะไฟล์ข้อมูลธรรมดา นั่นคือเก็บข้อมูลไว้ในไฟล์เพียงไฟล์เดียว เช่นเดียวกับไฟล์ข้อมูลของ Microsoft Access

SQLite เหมาะสำหรับแอปพลิเคชันแบบ Standalone แต่สามารถนำไปประยุกต์ ใช้งานได้หลากหลาย เช่น ดิกชันนารี

แคตล็อกสินค้า โปรแกรมแบบสอบถาม การเก็บข้อมูลที่ต้องการส่งเป็นไฟล์ข้อมูลทางเมลหรือโทรศัพท์มือถือ เป็นต้น

2.3 การดูแลสุนัข

ลักษณะทั่วไปของสุนัขชิสุ [3] เป็นสุนัขขนาดเล็ก มีน้ำหนักประมาณ 4.5 - 7.5 กิโลกรัม ส่วนสูงประมาณ 25 - 27 เซนติเมตร ชิสุมีลักษณะนิสัยขี้กลัวหาญ มีความตื่นตัว ขี้ประจบ มีความสง่าอยู่ในตัว นอกจากนี้ ชิสุ ยังรักความสะอาด เป็นมิตรกับทุกคน ปรับตัวได้ดี และชอบที่จะมีส่วนร่วมในกิจกรรมต่างๆ กับเจ้าของ

ลักษณะทั่วไปของสุนัขปอมเมอเรเนียน [4] เป็นสุนัขที่มีขนาดกะทัดรัด น้ำหนักเฉลี่ยอยู่ระหว่าง 1.7-2.5 กิโลกรัม ปอมเมอเรเนียนเป็นสุนัขที่ว่องไวปราดเปรียว หางสวยงามเป็นพวง ลักษณะนิสัยของปอมเมอเรเนียน คือ จะตื่นตัวเสมอ เหา่แก่ง มีนิสัยขี้ขลาดขี้กลัวเห็น สง่างาม และขณะก้าวอย่างแสดงถึงความมีชีวิตชีวา เป็นพันธุ์สมบูรณ์ทั้งรูปร่างและการเคลื่อนไหว

ลักษณะทั่วไปของสุนัขบางแก้ว [5] จะมีขนาดเท่าสุนัขไทยหรือเล็กกว่าเล็กน้อย ขนฟูยาว สง่างาม ว่องไว แข็งแรง เวลาขึ้นมักเชิดหน้า หัวกะโหลกใหญ่ ปากยาวแหลม หูเล็กสั้นตั้งป้องไปข้างหน้า คอคล้ายจิ้งจอก ตาเล็กกลมรี จมูกสีดำ ฟันซี่เล็กขาวคม หางเป็นพวงสืบทอดมาจากสุนัขจิ้งจอก

ลักษณะทั่วไปของสุนัขพุดเดิ้ล [6] ถูกจัดอยู่กลุ่มสุนัขประเภทสวยงาม ปากเรียวยาว ดวงตากลมโต หูห้อยลงมาปิดแก้ว ขนดกและหยิกชนิดติดหนัง ขนสั้นและเงางาม ขนค่อนข้างละเอียด เรียบ หยิบเล็กน้อยและไม่มีขนปุกปุย สีขนมีตั้งแต่สีน้ำตาลอ่อนจนถึงน้ำตาลแก่

ลักษณะทั่วไปของสุนัขชิวาว่า [7] จัดได้ว่าเป็นสุนัขพันธุ์ที่มีขนาดเล็กที่สุดในโลก และในปัจจุบัน พันธุ์ชิวาว่ากำลังเป็นที่นิยมกันมากของเหล่าผู้ที่รักสุนัข เนื่องด้วยขนาดตัวที่เล็กกะทัดรัด ผู้คนส่วนใหญ่จึงนิยมนำชิวาว่ามาเลี้ยง

2.4 โรคที่มักพบในสุนัข

โรคไข้หัด [8] เกิดจากเชื้อไวรัส มักเกิดกับลูกสุนัขอายุตั้งแต่ 2-3 เดือน โดยอาการจะแสดงออกมาทางประสาทสามารถป้องกันได้โดยฉีดวัคซีน ตั้งแต่อายุ 2 เดือน เป็นเข็มแรก

หลังจากนั้นอีกหนึ่งเดือนก็พาไปรับการฉีดวัคซีนเข็มที่สองเป็นการกระตุ้นภูมิคุ้มกันและฉีดซ้ำทุก ๆ ปี ปีละ 1 ครั้ง

โรคพาร์โวไวรัสหรือลำไส้อักเสบ เป็นโรคที่มีการระบาดทั่วโลก สามารถเกิดการระบาดได้ง่ายรวดเร็วและรุนแรง สุนัขจะตายเนื่องจากเกิดท้องร่วงอย่างรุนแรง ร่างกายสูญเสียน้ำมาก โรคนี้พบมากในสุนัขอายุ 2-6 เดือน วิธีป้องกันคือ การฉีดวัคซีนตั้งแต่สุนัขอายุได้ 2 เดือน และกระตุ้นภูมิคุ้มกันโดยฉีดวัคซีนอีกครั้งเมื่ออายุได้ 3 เดือน หลังจากนั้นก็ฉีดกระตุ้นทุกปีปีละ 1 ครั้ง

โรคพิษสุนัขบ้า เกิดจากเชื้อไวรัสชื่อเรบีส์ (Rabies) ไวรัสนี้มีกักอาศัยอยู่ในระบบประสาท จะทำให้สุนัขมีอาการทางประสาท นอกจากสุนัขแล้ว เชื้อไวรัสนี้ยังสามารถติดต่อกันและสัตว์อื่นได้ ถ้าโดนสุนัขที่มีเชื้อกัด

พยาธิไส้เดือน อาศัยอยู่ในลำไส้สุนัข ขนาดยาวประมาณ 2-3 นิ้ว พบมากในลูกสุนัข การติดต่อของพยาธิชนิดนี้สามารถติดต่อได้โดยสุนัขกินไข่พยาธิเข้าไป หรือถูกขนไซผ่านทางผิวหนัง การป้องกันทำได้ โดยถ่ายพยาธิ หลังจากเกิดประมาณ 2-3 อาทิตย์ ก็ให้ถ่ายพยาธิทุกๆ 2-3 เดือน

ตารางที่ 1: ข้อมูลช่วงอายุและโปรแกรมการฉีดวัคซีน

อายุ	โปรแกรม
4-6 สัปดาห์	ถ่ายพยาธิและตรวจสุขภาพ
6-8 สัปดาห์	วัคซีนรวม 5 โรค (โรคไข้หัด ลำไส้อักเสบ ตับอักเสบ เลปโตสไปโรซีสและหลอดลมอักเสบ) ครั้งที่ 1 และ ถ่ายพยาธิ
10 สัปดาห์	วัคซีนรวม 5 โรค ครั้งที่ 2 และถ่ายพยาธิ
12 สัปดาห์	วัคซีนรวม 5 โรค ครั้งที่ 3 และวัคซีนโรคพิษสุนัขบ้า ครั้งที่ 1 ถ่ายพยาธิ
14 สัปดาห์	เริ่มป้องกันพยาธิหนอนหัวใจ
24 สัปดาห์	วัคซีนโรคพิษสุนัขบ้า ครั้งที่ 2 และถ่ายพยาธิ

หมายเหตุ : หลังจากนั้นจะต้องทำการฉีดวัคซีนรวมและวัคซีนโรคพิษสุนัขบ้าปีละ 1 ครั้ง ไปจนตลอดชีวิตและถ่ายพยาธิทุก 3 เดือน

2.5 โปรแกรมวัคซีนในสุนัข

ช่วงเวลาที่เหมาะสมสำหรับการฉีดวัคซีน [9] คือ ช่วงที่ลูกสุนัขมีระดับภูมิคุ้มกันจากแม่ลดลง จนระดับภูมิคุ้มกันเดิมในร่างกาย ไม่สามารถไปรบกวนการทำงานของวัคซีน ดังนั้นควรให้ลูกสุนัข ได้รับการฉีดวัคซีนเมื่ออายุประมาณ 6-8 สัปดาห์

2.6 งานวิจัยที่เกี่ยวข้อง

Dogsly แอปพลิเคชันแชร์รูปสำหรับคนรักสุนัข [10] เป็นแอปพลิเคชันฟรีบนระบบปฏิบัติการ iOS สำหรับแชร์รูปถ่ายของสุนัขโดยเฉพาะ ไม่ว่าจะเลี้ยงสุนัขไว้เอง หรือไปเจอสุนัขน่ารักที่ไหน ก็สามารถถ่ายรูปแล้วแบ่งปัน หรือกดแชร์รูปขึ้นไปยังทวิตเตอร์ได้ นอกจากนี้ยังเป็นที่รวมตัวของเหล่าผู้รักสุนัข และใช้เป็นช่องทางสำหรับแลกเปลี่ยนซื้อขายสุนัขอีกด้วย

Pintofeed แอปพลิเคชันให้อาหารแก่สุนัข [11] พัฒนาโดย Carlos Herrera แอปพลิเคชันนี้ควบคุมผ่านทางสมาร์ตโฟนของเจ้าของสุนัขและเครื่องให้อาหารไร้สาย โดยสามารถตรวจสอบเวลาการกินอาหาร และจำนวนอาหารที่เหลือ โดยจะส่งข้อมูลรายงานเข้ามายังข้อความ อีเมล ทวิตเตอร์ของเจ้าของราคาประมาณ 4,600 บาท

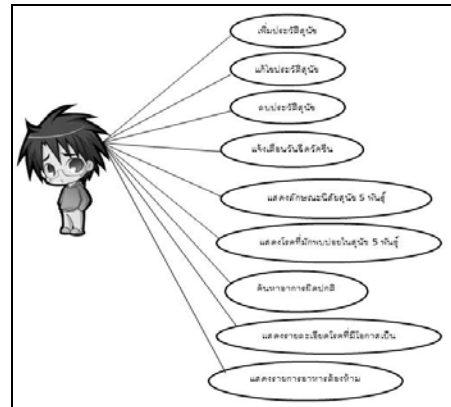
3. วิธีการดำเนินการศึกษาวิจัย

1. ศึกษา ค้นคว้า แนวทาง ทฤษฎี เอกสารต่าง ๆ ของงานที่เกี่ยวข้อง
2. รวบรวมข้อมูลการฉีดวัคซีนและการดูแลสุนัขในสายพันธุ์ต่างๆ
3. วิเคราะห์ออกแบบระบบงานของแอปพลิเคชัน
4. เขียนโปรแกรมของแอปพลิเคชันด้วยภาษาจาวา
5. ทดสอบระบบและแก้ไขข้อบกพร่องของแอปพลิเคชัน
6. ประเมินผลการทำงานของแอปพลิเคชัน โดยนำไปทดสอบการใช้งาน กับผู้ใช้ที่เกี่ยวข้อง
7. จัดทำเอกสารคู่มือการใช้งานและสรุปผลการดำเนินงาน

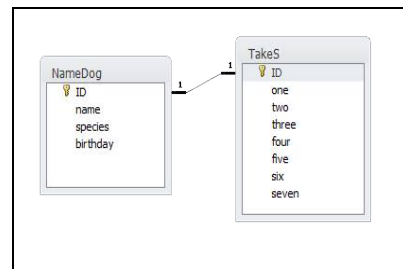
4. ผลการดำเนินงาน

4.1 ผลการวิเคราะห์และออกแบบแอปพลิเคชัน

โดยอาศัย Use case Diagram มาวิเคราะห์และออกแบบระบบ เพื่ออธิบายการทำงานของระบบงานเป็นภาพประกอบ และออกแบบตารางประวัติสุนัข เพื่อใช้คำนวณช่วงอายุที่ต้องพาไปฉีดวัคซีน และตารางกำหนดการฉีดวัคซีนแต่ละเข็ม เพื่อจัดเก็บลงในฐานข้อมูล SQLite ดังตารางที่ 2 – 3



ภาพที่ 1: Use Case Diagram ภาพรวมของระบบ



ภาพที่ 2: ความสัมพันธ์ระหว่างตาราง NameDog และ TakeS

ตารางที่ 2: ข้อมูลในตาราง NameDog

Table name : NameDog				
Description : ประวัติสุนัข				
Primary Key : ID				
Attribute	Type	Size	Description	Constraint
ID	Integer		ลำดับสุนัข	PK
name	Text	100	ชื่อสุนัข	
species	Text	100	สายพันธุ์	
birthday	Text	100	วันเกิด	

ตารางที่ 3: ข้อมูลในตาราง TakeS

Table name : TakeS				
Description : กำหนดฉีดวัคซีน				
Primary Key : ID				
Attribute	Type	Size	Description	Constraint
ID	Integer		ลำดับสุนัข	PK
one	Text	5	วัคซีนเข็มที่ 1	
two	Text	5	วัคซีนเข็มที่ 2	
three	Text	5	วัคซีนเข็มที่ 3	
four	Text	5	วัคซีนเข็มที่ 4	
five	Text	5	วัคซีนเข็มที่ 5	
six	Text	5	วัคซีนเข็มที่ 6	
seven	Text	5	วัคซีนเข็มที่ 7	

4.2 ผลการพัฒนาแอปพลิเคชัน

จากการพัฒนาแอปพลิเคชัน การดูแลสุขภาพของสุนัข สามารถแสดงรายละเอียดแอปพลิเคชัน ออกเป็น 4 ส่วนหลัก ดังแสดงในภาพที่ 3 - 11 ดังนี้



ภาพที่ 3: หน้าจอหลักของแอปพลิเคชัน



ภาพที่ 4: หน้าจอกรอกประวัติสุนัข



ภาพที่ 5: หน้าจอการแจ้งเตือนฉีดวัคซีน



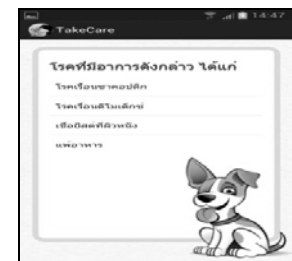
ภาพที่ 6: หน้าจอการดูแลสุนัขสายพันธุ์ต่างๆ



ภาพที่ 7: หน้าจอเลือกสายพันธุ์ของสุนัข



ภาพที่ 8: หน้าจอค้นหาโรคของสุนัข



ภาพที่ 9: หน้าจอโรคของสุนัขตามอาการที่เลือก



ภาพที่ 10: หน้าจอรายการอาหารต้องห้าม



ภาพที่ 11: หน้าจอเมื่อกดที่รายการอาหารต้องห้าม

4.3 ผลการประเมิน

ผลการประเมินความพึงพอใจ การใช้งานโปรแกรมดูแลสุขภาพสุนัขบนโทรศัพท์มือถือ กับผู้ใช้งานที่เกี่ยวข้อง แบ่งเป็นสัตว์แพทย์ 3 คน คิดเป็นร้อยละ 9.09 และผู้ใช้งานทั่วไป 30 คน คิดเป็นร้อยละ 90.91 รวมทั้งสิ้นจำนวน 33 คน แบ่งเป็นผู้ชาย 18 คน คิดเป็นร้อยละ 54.55 เป็นผู้หญิง 15 คน คิดเป็นร้อยละ 45.45 ทำแบบสอบถาม และประเมินผลการใช้งานโปรแกรมพบว่า ผลการประเมินคิดเป็นระดับความพึงพอใจในภาพรวม ส่วนใหญ่มีความพึงพอใจอยู่ในระดับดี ได้ค่าคะแนนเฉลี่ยคิดเป็น 4.35 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.60 แสดงว่าผู้ใช้งานมีความพึงพอใจ โปรแกรมสามารถนำไปใช้งานได้ อยู่ในระดับดี

5. สรุป

จากการศึกษาและพัฒนาแอปพลิเคชันดูแลสุขภาพสุนัขบนโทรศัพท์มือถือ และได้ทำการทดสอบและประเมินผลการใช้งานกับผู้ใช้งานที่เกี่ยวข้อง พบว่าแอปพลิเคชันสามารถช่วยในการดูแลสุขภาพของผู้ใช้งาน ช่วยอำนวยความสะดวกให้กับผู้ใช้งาน และช่วยให้การเลี้ยงสุนัขมีประสิทธิภาพเพิ่มมากขึ้น

อย่างไรก็ตาม จากการทดสอบการใช้งาน มีข้อเสนอแนะเพิ่มเติมจากผู้ใช้งาน คือ การแจ้งเตือนในแต่ละครั้ง ยังไม่สามารถบอกรายละเอียดของวัคซีนแต่ละเข็มได้ และแอปพลิเคชันควรจะสามารถเพิ่มรูปสุนัขลงไปได้

เอกสารอ้างอิง

- [1] Jeck Ryo, การพัฒนาแอปพลิเคชันบนระบบปฏิบัติการแอนดรอยด์ สืบค้นเมื่อ 25 พฤศจิกายน 2555, จาก <http://android-developer-cpe.blogspot.com/2012/03/android.html>
- [2] Thaixox, SQLite สืบค้นเมื่อ 26 พฤศจิกายน 2555, จาก <http://my.thaixox.net/modules.php?name=News&file=article&sid=50>
- [3] กระจุกดอกทอคอม, ลักษณะทั่วไปของสุนัขพันธุ์ชิวสตี สืบค้นเมื่อ 7 ธันวาคม 2555, จาก <http://pet.kapook.com/view156.html>
- [4] กระจุกดอกทอคอม, ลักษณะทั่วไปของสุนัขพันธุ์ปอมเมอเรเนียน สืบค้นเมื่อ 7 ธันวาคม 2555, จาก <http://pet.kapook.com/view157.html>
- [5] กระจุกดอกทอคอม, ลักษณะทั่วไปของสุนัขพันธุ์บางแก้ว สืบค้นเมื่อ 7 ธันวาคม 2555, จาก <http://pet.kapook.com/view154.html>
- [6] กระจุกดอกทอคอม, ลักษณะทั่วไปของสุนัขพันธุ์ฟูเดิ้ล สืบค้นเมื่อ 7 ธันวาคม 2555, จาก <http://pet.kapook.com/view150.html>
- [7] กระจุกดอกทอคอม, ลักษณะทั่วไปของสุนัขพันธุ์จิรวาว สืบค้นเมื่อ 7 ธันวาคม 2555, จาก <http://pet.kapook.com/view1121.html>
- [8] Thailand Dogshow, 8 อันดับโรคของสุนัขที่ต้องระวัง สืบค้นเมื่อ 26 พฤศจิกายน 2555, จาก <http://www.thailanddogshow.com/content-8-อันดับโรคของสุนัขที่ต้องระวัง-4-127-3043-1.html>
- [9] Lovedog, โปรแกรมวัคซีนในสุนัข สืบค้นเมื่อ 26 พฤศจิกายน 2555, จาก <http://www.vs.mahidol.ac.th/lovedog/index.php/2012-01-27-03-03-04>
- [10] Dogslly, โซเชียลเน็ตเวิร์คของคนรักสุนัข สืบค้นเมื่อ 20 มกราคม 2556, จาก <http://it-in-thai.blogspot.com/2012/03/dogslly.html>
- [11] Pintofeed, แอปพลิเคชันใหม่เอาใจเจ้าของชอบลืมหิวอาหารสุนัข สืบค้นเมื่อ 13 มีนาคม 2557, จาก <http://www.dogilike.com/content/dailydog/1986/>

ระบบสารสนเทศงานสวัสดิการ กรณีศึกษา การไฟฟ้าฝ่ายผลิตแห่งประเทศไทย

Welfare Information System : Case Study of Electricity Generating Authority of Thailand

สมฤดี เจริญวารี (Somrudee Jaroenwaree)¹ และ มาลีรัตน์ โสदानิล (Maleerat Sodanil)²

¹ สาขาวิชาระบบสารสนเทศเพื่อการจัดการ ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ

² ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
mashimarotuk@gmail.com, msn@kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อนำเสนอระบบสารสนเทศสวัสดิการของการไฟฟ้าฝ่ายผลิตแห่งประเทศไทย (กฟผ.) ซึ่งพัฒนาในรูปแบบ web-based application โดยใช้ภาษา Active Server Page (ASP) และฐานข้อมูล SQL Server 2005 ระบบจะแสดงสารสนเทศที่เกี่ยวข้องกับสวัสดิการของพนักงาน เช่น ค่ารักษาพยาบาล ค่าเล่าเรียน ค่าธรรมเนียมการคลอดบุตร และผลประโยชน์อื่นๆ ที่เป็นส่วนหนึ่งของการจัดสวัสดิการของ กฟผ. ระบบสามารถแสดงข้อมูลทางสถิติในรูปแบบของกราฟและตารางที่ผู้บริหารสามารถนำข้อมูลไปใช้ในการวิเคราะห์และการวางแผนเกี่ยวกับงบประมาณสวัสดิการได้อย่างสะดวกและรวดเร็ว ผลจากการประเมินความพึงพอใจโดยผู้เชี่ยวชาญ 5 คน ได้ค่าเฉลี่ยเท่ากับ 4.46 และค่าความเบี่ยงเบนมาตรฐานเท่ากับ 0.60 ดังนั้นสรุปได้ว่า ระบบมีความพึงพอใจอยู่ในระดับดี

คำสำคัญ : ระบบสารสนเทศ สวัสดิการ

Abstract

The objective of this research was to present the information system according to welfare data of Electricity Generating Authority of Thailand (EGAT) that performed in a web-based application. The system was developed to using Active server Page (ASP)

language and SQL Server 2005 relating to employee welfare, such as medical expenses, tuition fees, maternity fees and other benefits that is as part of EGAT's welfare. This system can show the welfare information of employee in pictures, graphs and tables, for the executive's decision in Human Resource Management and Planning. The result of value mean from user's assessment is 4.46 and the value of standard deviation is 0.60. The result of evaluation indicated that the level of satisfaction of the system is good.

Keyword : Information System, Welfare

1. บทนำ

การไฟฟ้าฝ่ายผลิตแห่งประเทศไทย (กฟผ.) เป็นหน่วยงานรัฐวิสาหกิจภายใต้การกำกับดูแลของกระทรวงพลังงาน มีภารกิจหลักในการผลิตและส่งไฟฟ้าให้แก่การไฟฟ้านครหลวง การไฟฟ้าส่วนภูมิภาค และอุตสาหกรรมขนาดใหญ่ ซึ่งเป็นผู้ใช้ไฟฟ้าโดยตรง กฟผ. มีบุคลากรจำนวนมาก ทำให้นโยบาย กฟผ. ที่จะสนับสนุนระบบการจัดสวัสดิการให้แก่ครอบครัวของพนักงาน กฟผ. เพื่อช่วยให้พนักงาน กฟผ. มีคุณภาพชีวิตที่ดี มีขวัญและกำลังใจในการปฏิบัติหน้าที่อย่างมีประสิทธิภาพ

ข้อมูลงานสวัสดิการของพนักงานและครอบครัว การไฟฟ้าฝ่ายผลิตแห่งประเทศไทย เป็นสิ่งสำคัญต่อการบริหารงานขององค์กร เพื่อใช้เป็นข้อมูลประกอบการตัดสินใจของผู้บริหาร

ขององค์กร ที่นำข้อมูลสถิติค่าใช้จ่ายสวัสดิการไปปรับปรุงเงินสวัสดิการในปัจจุบันให้ดียิ่งขึ้น และใช้ข้อมูลในการบริหารจัดการในแต่ละปี เพื่อสำรองค่าใช้จ่ายขององค์กร แต่เดิมเมื่อผู้บริหารต้องการข้อมูลค่าใช้จ่ายเงินสวัสดิการ ต้องนำข้อมูลจากระบบ SAP ซึ่งในปัจจุบันระบบ SAP ยังไม่เอื้ออำนวยในการรายงานในรูปแบบกราฟ และไม่มีการสรุปให้เป็นสารสนเทศที่ตรงกับความต้องการใช้งานสำหรับผู้บริหารและผู้ที่เกี่ยวข้อง จึงต้องนำข้อมูลออกมาในรูปแบบ Excel การแสดงข้อมูลเป็นที่ยุ่งยากมากและใช้เวลานาน เพราะต้องการข้อมูลหลาย ๆ ด้านมาประกอบ เช่น ของรายงานค่าใช้จ่ายเงินสวัสดิการบิดา มารดา แล้วต้องการดูของคู่สมรสหรือบุตร เพราะใช้ประกอบการพิจารณา ทำให้การออกรายงานมีความยุ่งยาก ดังนั้นผู้พัฒนาจึงได้หาแนวทางที่จะเพิ่มประสิทธิภาพของระบบสารสนเทศให้ดีขึ้น จึงมีแนวคิดที่จะนำเอาเทคโนโลยีระบบ Web Application ซึ่งทำงานผ่านระบบเครือข่ายมาใช้ในการรวบรวมข้อมูลและแสดงผล เพื่อให้การจัดเก็บข้อมูลเป็นไปอย่างมีระบบ มีกระบวนการ มีวิธีการที่มีประสิทธิภาพ ผู้ใช้งานสามารถเข้าถึงระบบได้จากเครือข่ายภายในองค์กร

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ระบบสารสนเทศ

ระบบสารสนเทศ (Information system) หมายถึง ระบบที่ประกอบด้วยส่วนต่าง ๆ ได้แก่ ระบบคอมพิวเตอร์ทั้งฮาร์ดแวร์ ซอฟต์แวร์ ระบบเครือข่าย ฐานข้อมูล ผู้พัฒนาระบบ ผู้ใช้ระบบ พนักงานที่เกี่ยวข้อง ประมวลผลข้อมูลเพื่อสร้างสารสนเทศและส่งผลลัพธ์หรือสารสนเทศที่ได้ให้ผู้ใช้เพื่อช่วยสนับสนุนการทำงาน การตัดสินใจ การวางแผน การบริหาร การควบคุม การวิเคราะห์และติดตามผลการดำเนินงานขององค์กร

2.2 การจัดสวัสดิการและประเภทสวัสดิการ

หลักในการจัดสวัสดิการสำหรับครอบครัว มีความจำเป็นสำหรับสังคมไทยมาเป็นเวลานาน โดยเฉพาะครอบครัวที่มี

รายได้น้อย หรือครอบครัวที่ประสบปัญหาความเดือนร้อน ซึ่งกระทบต่อความเป็นอยู่ของสมาชิกในครอบครัว เพื่อบรรเทาปัญหา รวมทั้งส่งเสริม สนับสนุนให้ครอบครัวสามารถพึ่งพาตนเองได้อย่างยั่งยืน ซึ่งรูปแบบการจัดสวัสดิการสำหรับครอบครัวมีความหลากหลายและมีการพัฒนาโดยตลอด เพื่อให้เหมาะสมกับสภาพของสังคมไทยในปัจจุบัน มีทั้งลักษณะการช่วยเหลือเฉพาะหน้าและระยะยาว ด้วยการให้คำแนะนำ การเงิน สิ่งของและอำนวยความสะดวก โดยรัฐบาลตระหนักถึงความสำคัญของสถาบันครอบครัว จึงได้กำหนดนโยบายและแนวทางในการพัฒนาสถาบันครอบครัวไว้ในแผนทุกระดับ มีการออกกฎหมายและระเบียบต่าง ๆ รองรับการพัฒนาและจัดสวัสดิการสำหรับครอบครัว รวมทั้งกำหนดองค์กรหลักรับผิดชอบดำเนินการ

2.3 ภาษา HTML

HTML (Hypertext Markup Language) คือ ภาษาทางคอมพิวเตอร์ออกแบบมาเพื่อใช้ในการสร้างเว็บเพจ หรือข้อมูลอื่นที่เรียกดูผ่านทางเว็บเบราว์เซอร์ เริ่มพัฒนาโดย ทิม เบอร์เนิร์ส ลี (Tim Berners Lee) โดยใช้เครื่องมือในการเขียน Script ได้หลากหลาย เช่น Notepad, dreamweaver ซึ่ง HTML เป็น Script ซึ่งจะเรียงคำสั่งต่าง ๆ และเก็บไว้ในไฟล์ๆ หนึ่งที ผู้เขียนสร้างขึ้น จะแสดงผลผ่านทางเว็บเบราว์เซอร์ออกมาเป็นรูปแบบต่าง ๆ HTML เป็นมาตรฐานหนึ่งของ ISO ซึ่งจัดการโดย World Wide Web Consortium (W3C) และพัฒนารูปแบบของ HTML แบบใหม่ ที่เรียกว่า XHTML ซึ่งเป็นลักษณะของโครงสร้าง XML แบบหนึ่งที่มีหลักเกณฑ์ในการกำหนดโครงสร้างของโปรแกรมที่มีรูปแบบที่มาตรฐาน

2.4 IIS

IIS (Internet Information Service) เป็นโปรแกรมสำหรับการจำลองเครื่องของเราให้กลายเป็นเครื่องเว็บเซิร์ฟเวอร์ ซึ่งมีไว้ให้บริการด้าน Server ในรูปแบบต่าง ๆ ของ Internet เช่น Web Server, FTP Server, SMTP Server ฯลฯ ในระบบปฏิบัติการวินโดวส์ ถูกพัฒนาโดยบริษัทไมโครซอฟต์ ซึ่งในวินโดวส์เซิร์ฟเวอร์ 2003 นั้น เวอร์ชันของ IIS จะเป็นเวอร์ชัน 6.0 และนอกจากนี้ IIS 6.0 ยังได้รับการพัฒนาให้มี

ประสิทธิภาพการทำงานที่ดีขึ้น ทำให้สามารถรองรับการใช้งานต่าง ๆ ได้ดียิ่งขึ้นและล่าสุดบริษัทไมโครซอฟต์ ได้ออกเวอร์ชันใหม่นั้นคือ IIS 7.0

2.5 ภาษา ASP

ASP (Active Server Pages) เป็นเทคโนโลยีของไมโครซอฟต์ที่เพิ่มประสิทธิภาพของ HTML ให้สร้างเว็บเพจแบบไม่ตายตัว (Dynamic) เริ่มแรกมีเฉพาะบนเว็บเซิร์ฟเวอร์ IIS และ PWS บนระบบปฏิบัติการของไมโครซอฟต์ แต่ในปัจจุบันมีเครื่องมือเพื่อพัฒนาบน UNIX และ Windows NT ที่ไม่ได้ใช้ IIS หรือ PWS ได้แก่ Chile!Soft และ Instant ASP ของ Halcyonsoft สามารถใช้งานร่วมกับสคริปต์ที่เป็น VBScript หรือ JavaScript ได้เป็นอย่างดี รวมทั้งสามารถใช้ร่วมกับระบบฐานข้อมูลได้หลายรูปแบบ หลักการทำงานของ ASP จะทำงานโดยมีตัวแปรและเอ็กซีคิวต์ที่ฝั่งเซิร์ฟเวอร์ เรียกการทำงานว่าเป็น Server Side ส่วนการทำงานของบราวเซอร์ เราจะเรียกว่า Client Side การทำงานจะเริ่มจากที่ผู้ใช้ส่งความต้องการผ่านเว็บบราวเซอร์ทาง HTTP (HTTP Request)

2.6 ภาษา SQL

SQL (Structured Query Language) ระบบจัดการฐานข้อมูลที่ใช้ในปัจจุบันไม่ว่าจะเป็นเครื่องระบบใหญ่หรือไมโครคอมพิวเตอร์ สามารถใช้ SQL ที่ถือเป็นภาษาฐานข้อมูลที่มีความสามารถในการจัดการฐานข้อมูล ไม่ว่าจะเป็นในเรื่องของการสร้างฐานข้อมูล การเก็บ การเรียกใช้ หรือควบคุมข้อมูล คำสั่งใน SQL เป็นคำสั่งที่ง่ายและมีความสามารถที่จะใช้คำสั่งเพื่อที่จะเรียกข้อมูลจากฐานข้อมูลได้ทันที ประโยชน์ของ SQL ที่สามารถใช้ในการเรื่องของการนิยามข้อมูลการเรียกใช้หรือควบคุมคำสั่งเหล่านี้จะช่วยประหยัดเวลาในการพัฒนาระบบงานหรือนำไปใช้ในส่วนของการสร้างฟอร์ม การทำรายงานของระบบงานต่าง ๆ ได้รวดเร็วยิ่งขึ้น

2.7 งานวิจัยที่เกี่ยวข้อง

โชคชัย บุญมาลัย [1] ได้ศึกษาค้นคว้าเรื่อง ระบบคลังข้อมูลเพื่อการตัดสินใจของผู้บริหาร ธุรกิจการให้เช่าแบบลิสซิง เพื่อ

พัฒนาเครื่องมือสำหรับวิเคราะห์ข้อมูลและแสดงผลในรูปแบบรายงานหลายมิติ เพื่อเป็นประโยชน์ในการวางแผนและกำหนดทิศทางในการพัฒนาองค์กร ในการพัฒนาระบบคลังข้อมูลนี้ การพัฒนาเครื่องมือผลิตรายงานหลายมิติที่มีความยืดหยุ่นตามความต้องการของผู้ใช้(ผู้บริหาร) โดยใช้โปรแกรมดีที เดสทอปโปรแกรมพีชชานนอล พัฒนาส่วนต่อประสานกับผู้ใช้ จากผลการทดสอบด้วยการนำระบบไปให้ผู้บริหารทดลองใช้และผลจากการประเมินความพึงพอใจของผู้ดูแลระบบ และผู้บริหารด้วยแบบสอบถาม จำนวน 6 คน พบว่าผู้ดูแลระบบและผู้บริหารมีความพึงพอใจในระดับมาก สามารถใช้ระบบนี้เป็นเครื่องมือสำหรับวิเคราะห์ข้อมูลให้อยู่ในรูปแบบรายงานหลายมิติได้ และสามารถนำรายงานที่ได้ไปใช้ในการวางแผนและกำหนดทิศทางในการพัฒนาองค์กรได้อย่างมีประสิทธิภาพ

พิมพ์ใจ หารสกุล [2] ได้ศึกษาค้นคว้าเรื่อง เพื่อศึกษาความพึงพอใจของพนักงานการไฟฟ้าฝ่ายผลิตแห่งประเทศไทยที่มีต่อสวัสดิการที่การไฟฟ้าฝ่ายผลิตแห่งประเทศไทยจัดให้ครอบครัวของพนักงาน และเพื่อศึกษาข้อเสนอแนะการปรับปรุงสวัสดิการที่การไฟฟ้าฝ่ายผลิตแห่งประเทศไทยจัดให้แก่ครอบครัวของพนักงาน โดยการเลือกกลุ่มตัวอย่างแบบโควต้าและเก็บรวบรวมข้อมูลโดยใช้แบบสอบถามที่ผู้วิจัยสร้างขึ้น สถิติที่ใช้วิเคราะห์ข้อมูลได้แก่ ความถี่ ร้อยละ ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน ผลการวิจัยพบว่า พนักงาน กฟผ. ที่เป็นกลุ่มตัวอย่างมีความพึงพอใจต่อสวัสดิการที่ กฟผ. จัดให้ครอบครัวของพนักงานในระดับดี

ดารณี คงเอียด [3] ได้ค้นคว้าแบบอิสระเรื่อง คุณภาพการให้บริการด้านสวัสดิการต่อบุคลากร ภูมิศึกษา มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี เพื่อศึกษาคุณภาพการให้บริการด้านสวัสดิการต่อบุคลากร กลุ่มตัวอย่างที่ใช้ในการศึกษาจำนวน 300 ตัวอย่าง โดยใช้แบบสอบถามเป็นเครื่องมือในการวิจัย ส่วนสถิติที่ใช้ในการวิเคราะห์ข้อมูลได้แก่ ค่าความถี่ ร้อยละ ค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน ค่า Independent Samples t-test ใช้เมื่อมีข้อมูล 2 กลุ่ม ส่วนค่า One-Way ANOVA (F-test) ใช้เมื่อมีข้อมูลมากกว่า 2 กลุ่มขึ้นไป ถ้าพบว่ามีความแตกต่างจะเปรียบเทียบเป็นรายคู่โดยใช้ LSD ผลการวิเคราะห์

ข้อมูลด้านบุคลากร พบว่าส่วนมากเป็นเพศหญิง มีอายุตั้งแต่ 36-40 ปี ระดับการศึกษาปริญญาตรี สถานภาพสมรส มีอายุงานมากกว่า 15 ปีขึ้นไป รายได้ต่อเดือนต่ำกว่า 15,000 บาท และมีสถานะ/ตำแหน่งงานเป็นข้าราชการ ความคิดเห็นของบุคลากรเกี่ยวกับปัจจัยคุณภาพการบริการในภาพรวมและรายด้านพบว่ามีความคิดเห็นอยู่ในระดับเห็นด้วยปานกลาง

ปิยะวัฒน์ ทองแก้วและสมชาย ปรการเจริญ [4] ได้ทำวิจัยเรื่อง การพัฒนาระบบจัดเก็บและค้นคืนสารนิพนธ์ทางเทคโนโลยีสารสนเทศโดยใช้ศัพท์ควบคุมที่มีโครงสร้างแบบชิฮอร์สบนแนวคิดออนโทโลยีสำหรับห้องสมุดดิจิทัล ในลักษณะของเว็บแอปพลิเคชันซึ่งพัฒนาด้วยโปรแกรม PHP, MySQL เป็นฐานข้อมูล และทำการประเมินประสิทธิภาพการทำงานของระบบ โดยการใช้ค่าความแม่นยำและค่าความถูกต้อง ซึ่งผลการประเมินจากผู้เชี่ยวชาญจำนวน 5 คน พบว่าผลความพึงพอใจของโปรแกรมอยู่ในระดับดี เนื่องจากระบบสามารถทำงานได้ตรงตามความต้องการของผู้เชี่ยวชาญและแสดงผลลัพธ์ได้ตามที่ต้องการ

3. วิธีการดำเนินงาน

วิธีการดำเนินงานของระบบสารสนเทศงานสวัสดิการกรณีศึกษาการไฟฟ้าฝ่ายผลิตแห่งประเทศไทย ผู้พัฒนาได้แบ่งวิธีการดำเนินงานออกเป็น

3.1 ศึกษาข้อมูลและวิเคราะห์ความต้องการของระบบ

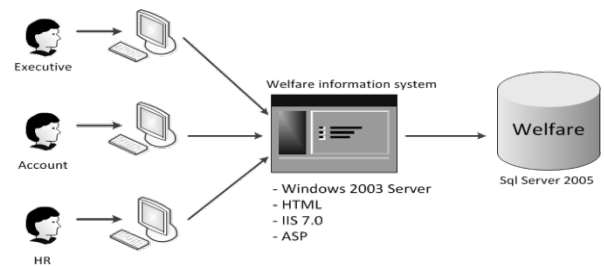
จากการศึกษาข้อมูล พบปัญหาในการจัดทำเป็นรายงาน ซึ่งมีขั้นตอนยุ่งยากและใช้เวลานาน ผู้พัฒนามีแนวคิดที่จะนำข้อมูลสวัสดิการมาทำเป็นสารสนเทศ โดยพัฒนาเป็นระบบเพื่อให้ผู้บริหารและผู้ที่เกี่ยวข้องได้รับประโยชน์จากการพัฒนา โดยในขั้นตอนการศึกษาข้อมูลและวิเคราะห์ปัญหานั้น ผู้พัฒนาได้รวบรวมเอกสารและรายงานที่เกี่ยวข้องกับงานสวัสดิการ พร้อมทั้งรวบรวมความต้องการในการใช้งานระบบจากผู้บริหารและผู้ที่เกี่ยวข้อง โดยการสัมภาษณ์และเก็บตัวอย่าง นำมาวิเคราะห์และออกแบบระบบต่อไป

3.2 ศึกษาเครื่องมือที่ใช้พัฒนาระบบ

ในการพัฒนาระบบสารสนเทศงานสวัสดิการ ผู้พัฒนาได้ออกแบบระบบให้เป็นรูปแบบของ web-based application โดยใช้ภาษา Active Server Page (ASP) เป็นเครื่องมือในการพัฒนา และใช้ SQL Server 2005 ในการจัดการระบบฐานข้อมูล

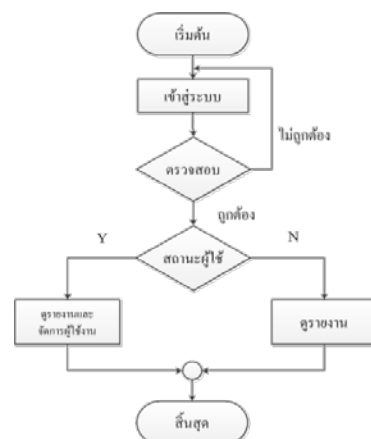
3.3 การวิเคราะห์และออกแบบระบบ

จากการวิเคราะห์ข้อมูลผู้พัฒนาได้ออกแบบระบบให้มีลักษณะ web-based application โดยมีสถาปัตยกรรมของระบบแบบ Client/Server ได้แสดงดังภาพที่ 1

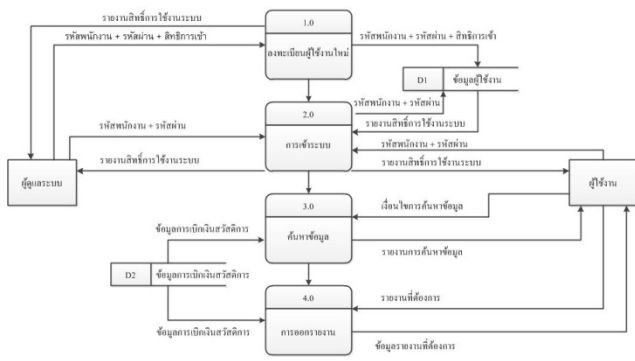


ภาพที่ 1: สถาปัตยกรรมของระบบแบบ Client/Server

เพื่อให้สามารถเข้าใจในหลักการ การทำงานโดยรวมของระบบง่ายขึ้น ผู้พัฒนาได้ใช้ Flow Chart Diagram ดังภาพที่ 2 และ Data Flow Diagram ในการวิเคราะห์ระบบ ซึ่งสามารถอธิบายด้วยผังการไหลของข้อมูล แสดงดังภาพที่ 3



ภาพที่ 2: ขั้นตอนการทำงาน



ภาพที่ 3: ผังการไหลของข้อมูลในระบบ

จากภาพที่ 3 ระบบประกอบด้วยงานหลักทั้งหมด 4 กระบวนการ ซึ่งสามารถอธิบายรายละเอียดได้ดังต่อไปนี้

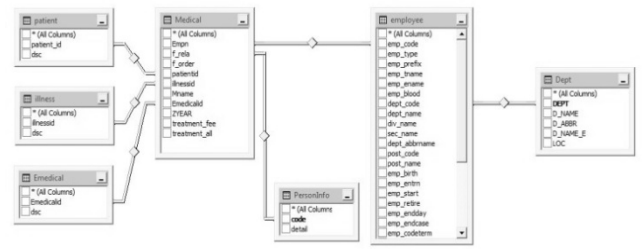
กระบวนการที่ 1.0 การลงทะเบียนผู้ใช้งานใหม่ เป็นกระบวนการที่ผู้ดูแลระบบจะต้องนำข้อมูลผู้ใช้งานใหม่เข้าสู่ระบบ เช่น รหัสพนักงาน รหัสผ่านและสิทธิการเข้า เพื่อทำการบันทึกข้อมูลลงในตารางข้อมูลผู้ใช้งาน

กระบวนการที่ 2.0 การเข้าสู่ระบบ เป็นกระบวนการตรวจสอบเช็คข้อมูลผู้ใช้และรหัสผ่านในการเข้าสู่ระบบของผู้ที่มีสิทธิ์ในการเข้าใช้ระบบ โดยผู้ใช้ใส่รหัสพนักงานและรหัสผ่าน

กระบวนการที่ 3.0 การค้นหาข้อมูล เป็นส่วนที่แสดงข้อมูลที่ผู้ใช้ต้องการค้นหา เช่น ค่ารักษาพยาบาล มีข้อมูลในด้านจำนวนผู้เบิกและจำนวนค่ารักษาพยาบาล มีการเลือกปีและเลือกสาขางาน

กระบวนการที่ 4.0 การออกรายงาน เป็นส่วนที่ผู้ใช้งานสามารถพิมพ์ข้อมูลการเบิกสวัสดิการเป็นรายงาน

ผู้พัฒนาได้ออกแบบระบบฐานข้อมูลขึ้นมาใหม่ โดยใช้ Entity Relationship Diagram เพื่อให้เห็นโครงสร้างและการเชื่อมโยงของข้อมูล ผู้พัฒนาได้ยกตัวอย่างตารางเก็บข้อมูลและความสัมพันธ์ของตารางที่เกี่ยวข้องในเรื่องการเบิกค่ารักษาพยาบาล ดังภาพที่ 4



ภาพที่ 4: แผนภาพความสัมพันธ์ของข้อมูล ส่วนการเบิกค่ารักษาพยาบาล

3.4 การพัฒนาระบบงาน

การพัฒนาระบบสารสนเทศงานสวัสดิการนี้ถูกสร้างและพัฒนามนระบบปฏิบัติการ Windows 2003 Server ใช้ Internet Information Server 7.0 (IIS) เป็นเว็บเซิร์ฟเวอร์ ส่วนของเว็บเพจถูกสร้างและพัฒนาด้วยภาษา ASP และ HTML โดยโปรแกรม Editplus 3.0 และใช้โปรแกรม Microsoft SQL Server 2005 เป็นระบบจัดการฐานข้อมูล

3.5 การทดสอบระบบสารสนเทศงานสวัสดิการ

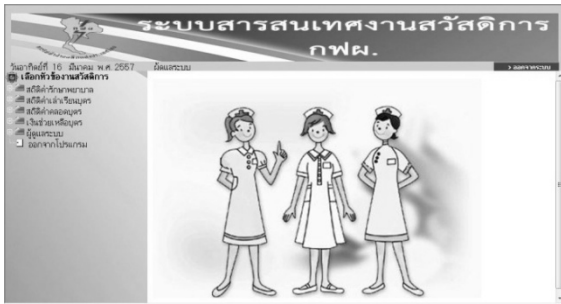
ผู้พัฒนาได้กำหนดการทดสอบออกเป็น 3 ขั้นตอน ได้แก่ การทดสอบในขั้นแอลฟา (Alpha Testing) การทดสอบในขั้นเบต้า (Beta Testing) และการประเมินความพึงพอใจโดยกลุ่มผู้เชี่ยวชาญ

การประเมินผลระบบสามารถแบ่งออกเป็น การประเมินความสามารถของระบบว่าตรงกับความต้องการของผู้ใช้ การประเมินความสามารถของระบบด้านการติดต่อผู้ใช้โปรแกรม การประเมินความถูกต้องในการทำงานของระบบ การประเมินระบบในด้านประสิทธิภาพว่าตรงตามที่ต้องการ และการประเมินระบบในด้านการรักษาความปลอดภัยของข้อมูล กลุ่มผู้เชี่ยวชาญทดสอบระบบจำนวน 5 คน

4. ผลการดำเนินงาน

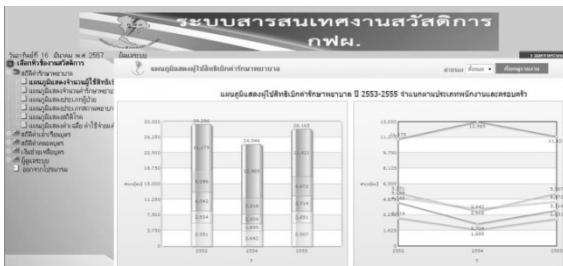
4.1 ผลการพัฒนาระบบ

ผลการพัฒนาระบบได้แบ่งหน้าจอใช้งานตามหน้าที่การทำงานเป็น 5 ส่วน คือ สติติค่ารักษาพยาบาล สติติค่าเล่าเรียน สติติค่าคลอดบุตร สติติเงินช่วยเหลือบุตรและผู้ใช้งาน



ภาพที่ 5: แสดงหน้าจอหลัก

ผลลัพธ์จากตัวอย่างหน้าจอแสดงข้อมูล



ภาพที่ 6: แสดงหน้าจอผลลัพธ์

4.2 ผลการประเมินความพึงพอใจ

ผลการประเมินความพึงพอใจของระบบในด้านความสามารถตามความต้องการผู้ใช้ มีค่าเฉลี่ยเท่ากับ 4.51:0.55 การติดต่อกับผู้ใช้งานระบบ มีค่าเฉลี่ยเท่ากับ 4.40:0.61 ความถูกต้องการทำงานของระบบ มีค่าเฉลี่ยเท่ากับ 4.60:0.52 ประสิทธิภาพของระบบ มีค่าเฉลี่ยเท่ากับ 4.30:0.60 ความปลอดภัยของระบบ มีค่าเฉลี่ยเท่ากับ 4.47:0.73 และผลการประเมินความพึงพอใจของระบบโดยมีค่าเฉลี่ยเท่ากับ 4.46:0.60

5. สรุป

จากการพัฒนาระบบสารสนเทศงานสวัสดิการนี้ขึ้นมา ทำให้การเพิ่มประสิทธิภาพของระบบสารสนเทศงานสวัสดิการดีขึ้น โดยได้พัฒนาระบบในรูปแบบ web-based application ซึ่ง

ทำงานผ่านระบบเครือข่ายมาใช้ในการรวบรวมข้อมูลและแสดงผล เพื่อให้การจัดเก็บข้อมูลเป็นไปอย่างมีระบบ มีกระบวนการ มีวิธีการที่มีประสิทธิภาพและสามารถนำมาใช้งานได้จริง จึงได้ทำการประเมินความพึงพอใจของระบบสารสนเทศงานสวัสดิการ โดยผู้เชี่ยวชาญ ซึ่งค่าเฉลี่ยเท่ากับ 4.46:0.60 สามารถสรุปผลการประเมินความพึงพอใจของระบบสารสนเทศงานสวัสดิการ กรณีศึกษา การไฟฟ้าฝ่ายผลิตแห่งประเทศไทยอยู่ในระดับดี

เอกสารอ้างอิง

- [1] โชคชัย บุญมาลัย, “ระบบคลังข้อมูลเพื่อการตัดสินใจของผู้บริหารธุรกิจการให้เช่าแบบลิสซิง”, มหาวิทยาลัยเชียงใหม่, 2551.
- [2] พิมพ์ใจ หารสกุล, “ความพึงพอใจของพนักงานการไฟฟ้าฝ่ายผลิตแห่งประเทศไทยที่มีต่อสวัสดิการที่การไฟฟ้าฝ่ายผลิตแห่งประเทศไทยจัดให้แก่ครอบครัวของพนักงาน”, มหาวิทยาลัยสุโขทัยธรรมราช, 2548.
- [3] ดารณี คงเอียด, “คุณภาพการให้บริการด้านสวัสดิการต่อบุคลากรกรณีศึกษา มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี”, มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี, 2554.
- [4] ปิยะวัฒน์ ทองแก้วและสมชาย ปราการเจริญ, “การพัฒนาระบบจัดเก็บและค้นหาสารนิพนธ์ทางเทคโนโลยีสารสนเทศ โดยใช้คำศัพท์ควบคุมที่มีโครงสร้างแบบซอร์สบนแนวความคิดออนโทโลยีสำหรับห้องสมุดดิจิทัล”, การประชุมทางวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 6 มิถุนายน 2553.
- [5] กิตติ ภัคดีวัฒนะกุลและไชยรัตน์ ปานปั้น, “ASP ฉบับฐานข้อมูล”, 2543.
- [6] พงษ์พันธ์ ศิวชัย, “SQL Server 2005 ฉบับสมบูรณ์”, 2549.

การพัฒนาระบบตรวจเยี่ยมสำหรับการประเมินตนเอง
โดยใช้โมเดลเซกิ กรณีศึกษามหาวิทยาลัยเนชั่น จังหวัดลำปาง
A Development of Peer Visit System for Self Assessment
by SECI Model : A Case Study of Nation University

บุรินทร์ รุจจนพันธุ์ (Burin Rujjanapan)
 คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเนชั่น
 thaiall@hotmail.com

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์ เพื่อพัฒนาระบบบันทึกผลการตรวจเยี่ยมคุณภาพการศึกษา ตามเกณฑ์ประกันคุณภาพการศึกษาของ สกอ. และสมศ. เพื่อพัฒนาระบบรายงานผลการตรวจเยี่ยมของกรรมการตรวจเยี่ยมในระดับคณะวิชา เป็นกรณีศึกษามหาวิทยาลัยเนชั่น มีกลุ่มตัวอย่างที่ใช้วิธีเลือกแบบเฉพาะเจาะจงจำนวน 5 กลุ่ม ประกอบด้วยผู้บริหาร และหน่วยงานที่รับผิดชอบการประกันคุณภาพการศึกษา ผู้รับผิดชอบองค์ประกอบคุณภาพระดับมหาวิทยาลัย บุคลากรในคณะวิชาที่รับผิดชอบองค์ประกอบ หน่วยงานที่รับผิดชอบตัวบ่งชี้ และผู้ประเมินภายใน เครื่องมือที่ใช้พัฒนาระบบคือเครื่องบริการอาปาเช่ ตัวแปลภาษาพีเอชพี เอแจ็กซ์ และระบบฐานข้อมูลมายเอสคิวแอล แล้วดำเนินการพัฒนาระบบฐานข้อมูลโดยใช้การจัดการความรู้ตามโมเดลเซกิ (SECI Model) ประกอบด้วย การแลกเปลี่ยนเรียนรู้ (Socialization) การสกัดความรู้ออกจากตัวคน (Externalization) การรวบรวมความรู้ (Combination) และการผนึกฝังความรู้ (Internalization) แล้วใช้แบบประเมินความพึงพอใจต่อการใช้งาน โปรแกรมต้นแบบ และภายหลังการใช้งานระบบโดยผู้ตรวจเยี่ยม งานที่พัฒนาขึ้นใช้กับการตรวจเยี่ยมระดับคณะวิชาตามรายงานการประกันคุณภาพการศึกษาปีการศึกษา 2555 โดยลักษณะคำตอบเป็นมาตราส่วนประมาณค่า

5 ระดับ พบว่า ผลประเมินความพึงพอใจต่ออบรมการใช้งานระบบฐานข้อมูลการตรวจเยี่ยมโดยรวมอยู่ระดับมาก ($\bar{x}=3.94$, $s.d.=0.46$) และผลการประเมินความพึงพอใจต่อการใช้งานระบบ อยู่ระดับมาก ($\bar{x}=3.65$, $s.d.=0.64$) ซึ่งสรุปได้ว่าระบบที่พัฒนาขึ้นตอบสนองความพึงพอใจของผู้ตรวจเยี่ยมได้
คำสำคัญ: การตรวจเยี่ยม ผู้ตรวจเยี่ยม การประกันคุณภาพ เกณฑ์ประเมินคุณภาพการศึกษา

Abstract

The objective of this operational research is to develop the peer visit system for the quality assessment system that is supported the internal and external indicators. It will help the inspector to generate the suggestion report. The system will be used as the case study of Nation University. The data sampling is divided into 5 groups, consisting of administrator, assessors, faculties, key performance indicator owner and peer visit committee. The system development tools are Apache Server, PHP Interpreter, AJAX and MySQL Database. Another tool for knowledge management process is SECI model that consist 4 sub-processes of socialization, externalization, combination and internalization. 5-scale rating questionnaire is collected to evaluate system performance in 2 times. This system was served for peer visit inspector in all faculties in 2012. The evaluation

result of training's satisfaction is high with a mean of 3.94 and standard deviation of 0.46. The evaluation result of system's satisfaction is high with a mean of 3.65 and standard deviation of 0.64. It is concluded that the system performance can satisfy the peer visit inspectors.

Keywords: Peer visit, Inspector, Quality Assurance, Quality Indicator

1. บทนำ

สถาบันการศึกษาในประเทศไทยต้องจัดการศึกษาให้มีคุณภาพมาตรฐาน เพื่อตอบสนองต่อเจตนารมณ์แห่งพระราชบัญญัติการศึกษาแห่งชาติ พ.ศ.2542 เป็นไปตามตัวบ่งชี้ที่จัดทำขึ้นโดยสำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (สมศ.) และสำนักงานคณะกรรมการการอุดมศึกษา (สกอ.) โดยผลการประเมินสะท้อนถึงคุณภาพตามเกณฑ์ที่กำหนดขึ้น

มหาวิทยาลัยในชั้นได้จัดทำรายงานการประเมินตนเองเพื่อรับการประเมินคุณภาพภายนอกประจำปีการศึกษา 2549 จากรายงานการประเมินคุณภาพภายนอก [1] ได้มีข้อเสนอแนะและทิศทางการพัฒนาเร่งด่วนของสถาบัน ว่ามหาวิทยาลัยควรพัฒนาฐานข้อมูล เกณฑ์ และเป้าหมาย เพื่อการประกันคุณภาพการศึกษาตามตัวบ่งชี้

ผลการประเมินในปีต่อมาถึงปัจจุบันยังต้องมีการพัฒนาให้สูงขึ้น การเพิ่มระบบและกลไกในการควบคุมคุณภาพจำเป็นต้องได้รับการสนับสนุนจากทุกระดับเชิงบูรณาการ การเพิ่มระบบตรวจเยี่ยมก่อนที่คณะวิชาจะจัดทำรายงานการประเมินตนเองฉบับสมบูรณ์ เป็นกลไกที่ทำให้เกิดการพัฒนาและมีการแลกเปลี่ยนเรียนรู้ การใช้เทคโนโลยีสารสนเทศเป็นสิ่งที่จะทำให้การไหลวนเป็นวงจรของข้อเสนอแนะและสารสนเทศมีประสิทธิภาพยิ่งขึ้น

จึงเป็นที่มาของการพัฒนาระบบตรวจเยี่ยม โดยมีวัตถุประสงค์เพื่อพัฒนาระบบบันทึกผลการตรวจเยี่ยมคุณภาพการศึกษา ตามเกณฑ์ประกันคุณภาพการศึกษาของ สกอ. และสมศ. เพื่อพัฒนาระบบรายงานผลการตรวจเยี่ยมของกรรมการ

ตรวจเยี่ยมในระดับคณะวิชา ซึ่งจะได้ระบบตรวจเยี่ยมที่สนับสนุนการจัดทำรายงานการประเมินตนเองฉบับสมบูรณ์อย่างเป็นระบบ

2. วรรณกรรมที่เกี่ยวข้อง

2.1 การประกันคุณภาพ

พระราชบัญญัติการศึกษาแห่งชาติ พ.ศ.2542 แก้ไขเพิ่มเติม (ฉบับที่ 2) พ.ศ.2545 ได้กำหนดจุดมุ่งหมายของการจัดการศึกษาที่มุ่งเน้นคุณภาพและมาตรฐาน โดยกำหนดรายละเอียดไว้ในหมวด 6 มาตรฐานและการประกันคุณภาพการศึกษา ซึ่งประกอบด้วย “ระบบการประกันคุณภาพการศึกษาภายใน” และ “ระบบการประกันคุณภาพการศึกษาภายนอก” เพื่อใช้เป็นกลไกในการพัฒนาระดับคุณภาพและมาตรฐานของสถาบันอุดมศึกษา [2]

การประกันคุณภาพการศึกษาภายใน เป็นการสร้างระบบและกลไกในการควบคุม ตรวจสอบและประเมินการดำเนินงานของสถานศึกษาให้เป็นไปตามนโยบาย เป้าหมาย และระดับคุณภาพตามมาตรฐานที่กำหนด โดยหน่วยงานต้นสังกัดและสถานศึกษากำหนดให้มีระบบการประกันคุณภาพภายในสถานศึกษา ให้ถือว่าการประกันคุณภาพภายในเป็นส่วนหนึ่งของกระบวนการบริหารการศึกษาที่ต้องดำเนินการอย่างต่อเนื่อง มีการจัดทำรายงานประจำปีที่เป็นรายงานการประเมินคุณภาพภายในเสนอต่อหน่วยงานต้นสังกัด และเปิดเผยต่อสาธารณชนเพื่อนำไปสู่การพัฒนาคุณภาพและมาตรฐานการศึกษาและรองรับการประกันคุณภาพการศึกษาภายนอก

การประเมินคุณภาพการศึกษาภายใน หมายถึง การดำเนินการตามระบบประเมินคุณภาพตามเกณฑ์ที่กำหนดโดยสำนักงานคณะกรรมการการอุดมศึกษา (สกอ.) ซึ่งเกณฑ์มุ่งเน้นการสะท้อนคุณภาพด้านระบบและกลไก และกระบวนการจัดการศึกษา ซึ่งประกอบด้วย 9 องค์ประกอบคุณภาพ

การประเมินคุณภาพการศึกษาภายนอก [3] หมายถึง การดำเนินการตามระบบประเมินคุณภาพตามเกณฑ์ที่กำหนดโดยสำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (สมศ.) ซึ่งเกณฑ์ต่าง ๆ มุ่งเน้นการสะท้อนคุณภาพด้านผลลัพธ์

จากระบบการจัดการศึกษา ซึ่งประกอบด้วย 18 มาตรฐาน และในปีการศึกษา 2555 ได้ถูกปรับให้มีภาระงานร่วมกับเกณฑ์ของ สกอ. ภายใต้กระบวนการที่ต้องดำเนินการประจำปี

2.2 โมเดล และเครื่องมือ

โมเดลเซกิ (SECI Model) [4] คือ แผนภาพแสดงความสัมพันธ์การหลอมรวมความรู้ในองค์กรระหว่างความรู้ฝังลึก (Tacit Knowledge) กับความรู้ชัดแจ้ง (Explicit Knowledge) ใน 4 กระบวนการ เพื่อยกระดับความรู้ให้สูงขึ้นอย่างต่อเนื่องเป็นวัฏจักร เริ่มจากการแลกเปลี่ยนเรียนรู้ (Socialization) การสกัดความรู้ออกจากตัวคน (Externalization) การรวบรวมความรู้ (Combination) และการผนึกฝังความรู้ (Internalization) และวนกลับมาเริ่มต้นทำซ้ำที่กระบวนการแรก เพื่อพัฒนาการจัดการความรู้ให้เป็นการประจำที่ยั่งยืน

ภาษาพีเอชพี (PHP Language) [5] คือ ภาษาคอมพิวเตอร์ประเภทโอเพนซอร์ส (Open Source Computer Language) สำหรับพัฒนาเว็บเพจแบบไดนามิก (Dynamic Webpage) เมื่อเครื่องบริการได้รับคำร้องจากผู้ใช้งานก็จะส่งให้กับตัวแปลภาษาทำหน้าที่ประมวลผลและส่งข้อมูลกลับไปยังเครื่องคอมพิวเตอร์ของผู้ใช้ที่ร้องขอ ในรูปของภาษาเอชทีเอ็มแอล ภาพ หรือแฟ้มอิเล็กทรอนิกส์อื่นใด ลักษณะของภาษามีรากฐานมาจากภาษาซีเป็นภาษาที่สามารถพัฒนาให้ใช้งานแบบโต้ตอบกับผู้ใช้ได้

2.3 งานวิจัยที่เกี่ยวข้อง

อัจฉรา แก้วละเอียด และผุสดี บุญรอด [6] ทำการศึกษาวิจัย เรื่อง การจัดการองค์ความรู้เชิงความหมายโดยใช้โมเดลเอสอีซีไอสำหรับตรวจสอบบัญชีคอมพิวเตอร์ เพื่อพัฒนาระบบการจัดการความรู้สำหรับตรวจสอบบัญชีคอมพิวเตอร์ กรณีศึกษา กรมตรวจบัญชีสหกรณ์ โดยสร้างฐานความรู้ออนโทโลยีในรูปแบบโดเมนออนโทโลยี แล้วแปลงไปเป็นภาษาเอ็กซ์เอ็มแอลรองรับการค้นคืนเชิงความหมาย ใช้เครื่องมือคือโปรแกรมโฮโซ ออนโทโลยี อีดิเตอร์ (Hozo-Ontology editor) มีวิธีการดำเนินงานวิจัย 3 ขั้นตอนคือ 1) การวิเคราะห์และออกแบบระบบจัดการองค์ความรู้โดยใช้โมเดลเอสอีซีไอ 2) การพัฒนาฐานความรู้และระบบการจัดการองค์ความรู้

3) การทดสอบระบบ โดยประเมินคุณภาพด้วยวิธีการแบบแบล็กบ็อกซ์ (Black box testing) มีผลประเมินโดยผู้เชี่ยวชาญในระดับดี ค่าเฉลี่ย 4.16 และส่วนเบี่ยงเบนมาตรฐาน 0.63 มีผลประเมินโดยผู้ใช้ในระดับดี ค่าเฉลี่ย 4.12 และส่วนเบี่ยงเบนมาตรฐาน 0.64 ซึ่งสรุปได้ว่าฐานความรู้ออนโทโลยีและระบบสารสนเทศที่พัฒนาขึ้นสามารถจัดการองค์ความรู้ได้อย่างเหมาะสม

ณัฐพล สมบูรณ์ และนลินภัทร์ ปรวัฒน์ปรีชกร [7] ทำการศึกษาวิจัย เรื่อง ระบบการจัดการความรู้ด้านเทคโนโลยีสารสนเทศ กรณีศึกษาสำนักงานปลัดกระทรวงการคลัง เพื่อพัฒนาระบบตามตอบด้านเทคโนโลยีสารสนเทศผ่านเว็บแอปพลิเคชัน โดยใช้เทคนิคการจัดการความรู้ด้วยโมเดลปลาทู (Tuna Model) ที่มี 3 ส่วนคือ 1) การรวบรวมและจัดเก็บความรู้แยกหมวดหมู่ด้านเทคโนโลยีสารสนเทศ (KV=Knowledge Vision) 2) การกระจายความรู้ (KS=Knowledge Sharing) 3) การเก็บเข้าคลังความรู้ (KA=Knowledge Assets) มีวิธีการดำเนินงานวิจัย 2 ขั้นตอนคือ 1) การวิเคราะห์และออกแบบระบบ 2) การพัฒนาระบบ มีผลประเมินความพึงพอใจโดยผู้เชี่ยวชาญในระดับดี ค่าเฉลี่ย 4.46 และส่วนเบี่ยงเบนมาตรฐาน 0.23 มีผลประเมินโดยผู้ใช้ในระดับดี ค่าเฉลี่ย 3.89 และส่วนเบี่ยงเบนมาตรฐาน 0.86 ซึ่งสรุปได้ว่าระบบสามารถนำไปประยุกต์ใช้งานได้จริงและช่วยในการถ่ายทอดความรู้ได้อย่างมีประสิทธิภาพ

กิตติยา สืออ่อน [8] ทำการศึกษาวิจัยเกี่ยวกับโมเดลเชิงสาเหตุประสิทธิผลการประกันคุณภาพภายในสำหรับกลุ่มสาขาวิทยาศาสตร์สุขภาพ โดยสังเคราะห์งานวิจัยทั้งในและต่างประเทศ พบว่า ปัจจัยที่ส่งผลต่อประสิทธิผลของการประกันคุณภาพการศึกษา ได้แก่ คุณลักษณะของบุคลากรด้านความตระหนักถึงความสำคัญและการมีทัศนคติที่ดีต่อการประกันคุณภาพภายใน การทำงานเป็นทีม การฝึกอบรมพัฒนาบุคลากร ความพร้อมด้านทรัพยากรทั้งด้านจำนวนบุคลากรงบประมาณ วัสดุอุปกรณ์และระบบเทคโนโลยีสารสนเทศ โครงสร้างองค์กรและการกำหนดให้มีผู้รับผิดชอบในการดำเนินงานประกันคุณภาพ วัฒนธรรมคุณภาพ ภาวะผู้นำของ

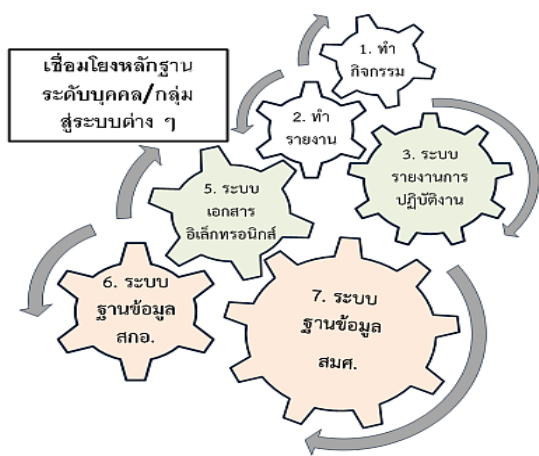
ผู้บริหาร และขนาดขององค์กร ส่วนประสิทธิผลการประกันคุณภาพการศึกษามี 2 ส่วนคือ ประสิทธิภาพระหว่างทาง ได้แก่ การบริหารจัดการแบบกระจายอำนาจ การจัดการเรียนการสอนที่เน้นผู้เรียนเป็นสำคัญ การปฏิบัติงานแบบวงจรคุณภาพ และความพึงพอใจในงานของบุคลากร สำหรับประสิทธิผลสุดท้าย ได้แก่ การบรรลุพันธกิจของสถาบัน

เมื่อศึกษาจากงานวิจัยที่เกี่ยวข้อง จึงได้เพิ่มกลไกการตรวจเยี่ยมและใช้โมเดลเชกซ์บ็อกซ์เคลื่อนที่กระบวนการในการพัฒนาระบบ เพื่อให้เกิดการจัดการความรู้ไปสนับสนุนการจัดทำรายงานการประเมินตนเองที่สมบูรณ์

3. วิธีการดำเนินการวิจัย

3.1 ศึกษาข้อมูลเบื้องต้น

3.1.1 รวบรวมข้อมูลจากเอกสาร และสัมภาษณ์ผู้เกี่ยวข้องทั้งทีมวิจัยและผู้มีบทบาทสำคัญต่องานประกันคุณภาพการศึกษา ทำความเข้าใจต่อกิจกรรม ข้อมูลและประเด็นปัญหา โดยมีเอกสารสำคัญคือ คู่มือการประกันคุณภาพการศึกษา และรายงานการประเมินตนเองที่ผ่านมา [1]



ภาพที่ 1 : การเชื่อมโยงหลักฐาน

3.1.2 ประชากร คือ อาจารย์และเจ้าหน้าที่จำนวน 88 คน มีกลุ่มตัวอย่างประกอบด้วย ผู้บริหาร และหน่วยงานที่รับผิดชอบการประกันคุณภาพการศึกษาจำนวน 7 คน ผู้รับผิดชอบองค์ประกอบคุณภาพระดับมหาวิทยาลัยจำนวน 8 คน บุคลากรในคณะวิชาที่รับผิดชอบองค์ประกอบจำนวน 10 คน หน่วยงานที่รับผิดชอบตัวบ่งชี้จำนวน 5 คน และผู้ประเมิน

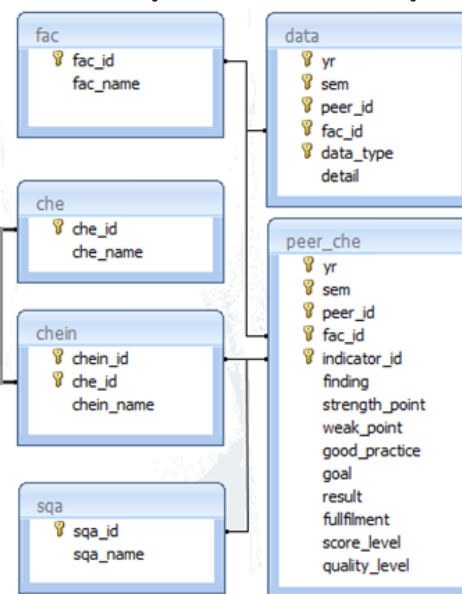
ภายใน 5 คน

3.2 กระบวนการจัดการความรู้

การพัฒนาระบบใช้กระบวนการจัดการความรู้ตามโมเดลเชกซ์ (SECI Model) ครอบคลุมวงจรการพัฒนา (SDLC) โดยเริ่มจากการแลกเปลี่ยนเรียนรู้ที่นำไปสู่การสร้างโปรแกรมต้นแบบ (Prototype) แล้วจัดเวทีอบรมให้ทดลองใช้ นำผลการใช้และข้อเสนอแนะไปปรับปรุงระบบ หลังมีการใช้งานระบบแล้ว คณะวิชาจะนำข้อมูลในระบบไปพัฒนาความรู้ภายในคณะวิชา และรับการประเมินจากผู้ประเมิน แล้วนำข้อเสนอแนะกลับไปจัดเวทีเรียนรู้ภายในคณะวิชาต่อไป

3.3 วิเคราะห์และออกแบบระบบงานใหม่

การวิเคราะห์ระบบใช้วิธีศึกษาจากคู่มือการประกันคุณภาพภายในและภายนอก โดยออกแบบตามทฤษฎีระบบฐานข้อมูลเชิงสัมพันธ์ เก็บข้อมูลในมายเอสคิวแอล พัฒนาโปรแกรมด้วยภาษาพีเอชพี และเทคนิคเอแจ็กซ์ ทำการพัฒนาโปรแกรม นำเสนอต่อผู้ตรวจเยี่ยม คณะวิชา และผู้เกี่ยวข้อง



ภาพที่ 2 : การบูรณาการเพิ่มอิเล็กทรอนิกส์

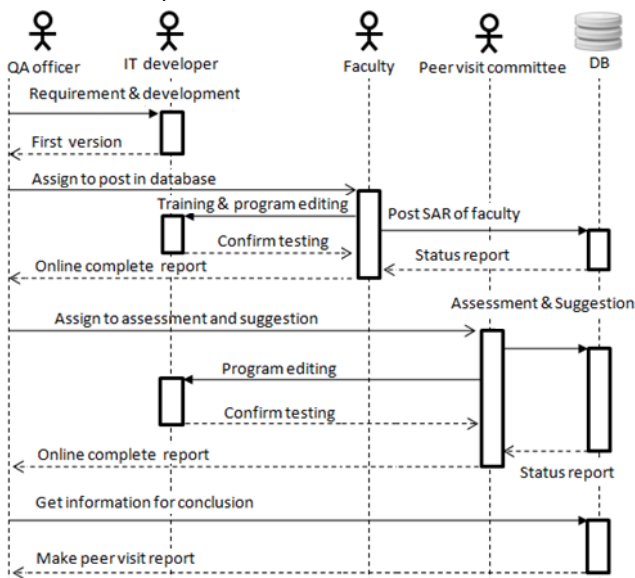
3.4 เก็บข้อมูลความพึงพอใจต่อการใช้งานโปรแกรม

เข้าเก็บข้อมูลจากคณะวิชา และผู้ประเมิน หลังผ่านช่วงเวลาของการประเมินคุณภาพการศึกษาระดับคณะวิชา และระดับมหาวิทยาลัย มีจำนวน 8 คน

3.4.1 ใช้ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐานแสดงระดับ

ความพึงพอใจ โดยแบบสอบถามใช้มาตราส่วนประมาณค่า (Rating Scale) แบบวัดเจตคติตามเทคนิคของลิเคิร์ท [9] 5 ระดับ คือ ดีมาก ดี พอใช้ น้อย และน้อยมาก

3.4.2 การแปลผลคะแนนสามารถสรุปเป็นระดับของความพึงพอใจ โดยนำระดับสูงสุดลบระดับต่ำสุดแล้วหารด้วยจำนวนช่วงที่ตั้งไว้ จึงได้เกณฑ์การแปลผลคะแนนดังนี้ มากที่สุดมีคะแนนเฉลี่ย 4.21–5.00 มากมีคะแนนเฉลี่ย 3.41–4.20 ปานกลางมีคะแนนเฉลี่ย 2.61–3.40 น้อยมีคะแนนเฉลี่ย 1.81–2.60 และน้อยที่สุดมีคะแนนเฉลี่ย 1.00–1.80



ภาพที่ 3 : ความสัมพันธ์ของกิจกรรมกับผู้รับผิดชอบ

4. ผลการดำเนินงาน

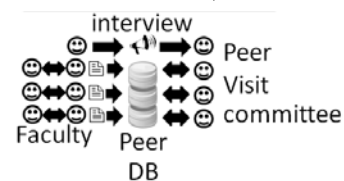
4.1 การวิเคราะห์สถานการณ์ระบบฐานข้อมูล พบว่ามหาวิทยาลัยควรมีระบบที่จะประกันได้ว่าเมื่อเข้ากระบวนการประเมินตนเองจะมีผลประเมินอยู่ในระดับที่สูงขึ้น จึงจัดให้มีการตรวจเยี่ยมก่อนจัดทำรายงานที่สมบูรณ์ เพื่อให้มีกลไกการแลกเปลี่ยนเรียนรู้ และปรับปรุงก่อนรับการประเมินจริงอย่างเป็นระบบ จึงได้มีการพัฒนาระบบบันทึกผลการตรวจเยี่ยมคุณภาพการศึกษา ตามเกณฑ์ประกันคุณภาพการศึกษาของ สกอ. และสมศ. และพัฒนาระบบรายงานผลการตรวจเยี่ยมของกรรมการตรวจเยี่ยมระดับคณะวิชา โดยดำเนินการพัฒนาระบบฐานข้อมูลที่บูรณาการกับการจัดการความรู้โดยใช้โมเดลเชก

4.2 การแลกเปลี่ยนเรียนรู้ (Socialization) เป็นการจัดเวทีระหว่างผู้ที่มีความรู้ฝังลึก (Tacit Knowledge) ที่เป็นผู้บริหารและทำงานด้านประกันคุณภาพมาร่วมแลกเปลี่ยนถึงบทเรียนจากการประเมินตนเองที่ผ่านมาและกำหนดประเด็นปัญหาที่ต้องการแก้ไข ได้ข้อสรุปเป็นการพัฒนาระบบตรวจเยี่ยม (Peer Visit) ที่จะยกระดับความสมบูรณ์ในการประเมินตนเอง



ภาพที่ 4 : กระบวนการทางสังคม

4.3 การสกัดความรู้ออกจากตัวคน (Externalization) เป็นการกำหนดบทบาท กลุ่มเป้าหมาย ทำความเข้าใจกิจกรรมกับผู้ที่มีความรู้ฝังลึกที่อยู่กับข้อมูลโดยตรง เพื่อให้สามารถสกัดและถ่ายทอดออกเป็นความรู้ชัดเจน โดยมีกิจกรรมอบรมเพื่อศึกษาความต้องการเพิ่มเติม กำหนดรูปแบบการพัฒนา ผ่านโปรแกรมต้นแบบที่นำเสนอในการอบรม พบว่า มีผลประเมินความพึงพอใจในระดับมาก ($\bar{x}=3.94, S.D.=0.40$)



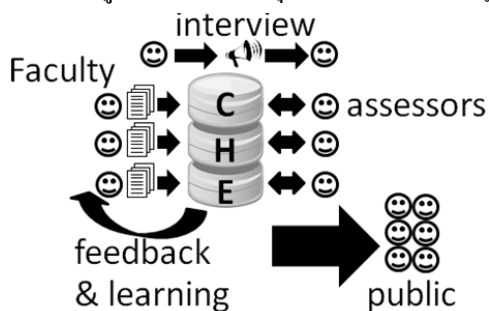
ภาพที่ 5 : กระบวนการสัมพันธ์ภายนอก

4.4 การรวบรวมความรู้ (Combination) เป็นการระบุนความรู้ชัดเจนที่เกิดจากการเข้าตรวจเยี่ยมของผู้รู้ และส่งข้อเสนอแนะเข้าไปในระบบให้คณะวิชาได้เข้าถึง แล้วนำผลในรูปความรู้ใหม่ไปแบ่งปัน วิพากษ์ แลกเปลี่ยนภายในคณะวิชาและผู้ตรวจเยี่ยมภายใน เกิดเป็นความเข้าใจร่วมเป็นความรู้ชัดเจนใหม่ นำไปสู่การจัดทำรายงานการประเมินตนเองที่สมบูรณ์ พบว่า มีผลประเมินความพึงพอใจของผู้ใช้งานระบบตรวจเยี่ยมอยู่ในระดับมาก ($\bar{x}=3.65, S.D.=0.64$)



ภาพที่ 6 : กระบวนการผสานองค์ความรู้

4.5 การผนึกฝังความรู้ (Internalization) เป็นการเลือก นำเสนอสารสนเทศในรูปแบบของรายงานสมบูรณแก่ผู้ประเมิน คุณภาพภายนอกในบทบาทของผู้ที่ตรวจประเมินความรู้ชัดแจ้ง ผ่านระบบฐานข้อมูลภายนอก (CheQA) ที่มีการเผยแพร่สู่ สาธารณะ แล้วให้ข้อเสนอแนะกลับแก่ผู้ถูกประเมิน ทำให้เกิด การพัฒนาเป็นความรู้ฝังลึกทั้งในคณะวิชา และมหาวิทยาลัย ส่งผลให้นำความรู้ไปพัฒนาปรับปรุงอย่างต่อเนื่องเป็นวัฏจักร



ภาพที่ 7 : กระบวนการส่งต่อเป็นความรู้ฝังลึก

5. สรุป

การใช้โมเดลเชกเป็นกลไกขับเคลื่อนการพัฒนาระบบ ตรวจสอบการประเมินตนเองสามารถใช้งานได้จริง ผลประเมิน ความพึงพอใจทั้งสองครั้งพบว่าอยู่ระดับมาก โดยครั้งแรกเป็น การฝึกอบรมและแลกเปลี่ยนระหว่างผู้รู้ด้านการประกัน คุณภาพ และผู้รู้ด้านเทคโนโลยีสารสนเทศ ครั้งที่สองเป็นการ ประเมินความพึงพอใจของผู้ตรวจเยี่ยมหลังสรุปผลการตรวจ เยี่ยม ด้วยแบบสอบถาม

แต่ผลประเมินความพึงพอใจครั้งแรกสูงกว่าครั้งที่สอง ซึ่ง เกิดจากความไม่พร้อมของคณะวิชาในการรับการตรวจเยี่ยม ตามร่างรายงานการประเมินตนเอง เมื่อผู้ประเมินเข้าตรวจเยี่ยม และได้รับข้อมูลที่ไม่สมบูรณ์ ทำให้ความพึงพอใจต่อระบบ ตรวจสอบลดลง เนื่องจากผู้ใช้มีความต้องการที่หลากหลาย และคาดหวังความสมบูรณ์ต่อข้อมูลที่คณะวิชาส่งเข้าไปใน ระบบ อนึ่งระบบนี้เป็นระบบที่ใช้งานจริงและมีนโยบายที่จะใช้ งานอย่างต่อเนื่อง จึงสามารถนำข้อเสนอแนะหลังการใช้งาน มา ปรับปรุงระบบให้พร้อมใช้งานสำหรับปีการศึกษาต่อไป เพื่อ สนับสนุนการพัฒนาการประกันคุณภาพการศึกษา

ข้อเสนอแนะ คือ ควรเพิ่มกระบวนการแลกเปลี่ยนเรียนรู้

จากต่างคณะวิชา และภายในคณะวิชาเดียวกัน โดยบริการระบบ ฐานข้อมูลตรวจเยี่ยมที่เชื่อมกับข้อเสนอแนะต่าง ๆ และเร่งการ จัดทำร่างรายงานการประเมินตนเองให้แล้วเสร็จโดยเร็ว ก่อน จัดให้มีการตรวจเยี่ยมอย่างเป็นระบบ และมีนโยบายกำกับให้ คณะวิชาดำเนินการตามแผนที่สามารถปฏิบัติได้จริง

เอกสารอ้างอิง

- [1] บุญรักษา สุนทรธรรม, “รายงานการประเมินคุณภาพภายนอก ระดับอุดมศึกษา มหาวิทยาลัยโยนก”, สำนักงานรับรองมาตรฐาน และประเมินคุณภาพการศึกษา, กรุงเทพฯ, 2550.
- [2] สำนักงานคณะกรรมการการอุดมศึกษา, “คู่มือการประกันคุณภาพ การศึกษาภายในสถานศึกษาระดับอุดมศึกษา”, กระทรวงศึกษาธิการ, กรุงเทพฯ, 2553.
- [3] สำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์กร มหาชน), “คู่มือการประเมินคุณภาพภายนอกรอบสาม” ระดับอุดมศึกษา, สำนักงานรับรองมาตรฐานและประเมินคุณภาพ การศึกษา (องค์กรมหาชน), กรุงเทพฯ, 2555.
- [4] Nonaka, I. and H. Takeuchi, “The Knowledge Creating Company”, Oxford University Press, New York, 1995.
- [5] พร้อมเลิศ หล่อวิจิตร, “คู่มือเรียน PHP และ MYSQL สำหรับผู้ เริ่มต้น”, บริษัทโปรวิชั่น จำกัด., กรุงเทพฯ, 2549.
- [6] อัจฉรา แก้วละเอียด และสุสติ บุญรอด, “การจัดการองค์ความรู้เชิง ความหมายโดยใช้โมเดลเอสอีซีไอสำหรับผู้ตรวจสอบบัญชี คอมพิวเตอร์,” the 9th National Conference on Computing and Information Technology (NCIT2013), กรุงเทพฯ, ประเทศไทย, 9-10 พ.ค.2556 หน้า 873-878.
- [7] ณัฐพล สมบูรณ์ และณลินภัทร์ ปรวัฒน์ปรีयर, “ระบบการจัดการ ความรู้ด้านเทคโนโลยีสารสนเทศ รมณศึกษา สำนักงานปลัด กระทรวงการคลัง,” the 9th National Conference on Computing and Information Technology (NCIT2013), กรุงเทพฯ, ประเทศไทย, 9-10 พ.ค.2556 หน้า 234-239.
- [8] กิตติยา ลีอ่อน, “โมเดลเชิงสาเหตุประสิทธิผลการประกันคุณภาพ ภายในสำหรับกลุ่มสาขาวิทยาศาสตร์สุขภาพ”, จุฬาลงกรณ์ มหาวิทยาลัย, กรุงเทพฯ, 2547.
- [9] L.W. Anderson, “ Likert Scales, Education Research Methodology and Measurement : An International Handbook”, John, D. Keeves, eds, Victoria : Pergamon pp.427-428, 1988.

การพัฒนา ระบบ BOTs เพื่อใช้สำหรับการแข่งขันเกมต่อสู้

Development of AI BOTs for the Fighting Game Competition

อังคณา สุวรรณจตุพร (Angkana Suwanjatuporn)¹ และภาณี อภิรมย์เสนีย์ (Paranee Sinsoongsud)²

¹ภาควิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยกรุงเทพ

²ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยกรุงเทพ

angkana.s@bu.ac.th, paranee.sins@bumail.net

บทคัดย่อ

เกมต่อสู้ (Fighting Game) เป็นเกมคอมพิวเตอร์ประเภทหนึ่งโดยลักษณะของเกมต่อสู้ (Fighting Game) นั้นจะมีผู้เล่น 2 ฝ่าย ในการสู้กันเพื่อหาผู้ชนะ ทั้งคู่จะทำหน้าที่ทั้งโจมตีและป้องกันการถูกโจมตีของฝ่ายตรงข้าม ยกตัวอย่างเช่น เกม Street Fighter ซึ่งเป็นเกมต่อสู้ที่ให้ผู้เล่นได้เลือกตัวละคร (Character) ที่ถูกบังคับโดยผู้เล่น เพื่อสู้กับ BOTs ที่ถูกกำหนดท่าทางที่จะนำมาใช้โดยผู้สร้าง ผลลัพธ์สุดท้ายเกมจะเลือกผู้ชนะที่มีพลังชีวิตเหลือมากที่สุด ด้วยเหตุนี้คณะผู้พัฒนาจึงได้เข้าร่วมการแข่งขัน Fighting Game AI Competition ในการพัฒนา AI BOTs ที่มีความฉลาดมากขึ้น เพื่อให้เป็นฝ่ายชนะ

ในการพัฒนาตัวละคร (Character) คณะผู้พัฒนาได้พัฒนาด้วยภาษาจาวา โดยใช้หลักการของ Artificial Intelligence (AI) ด้วยการใช้เทคนิคในการโจมตีอย่างต่อเนื่องโดยไม่เปิดโอกาสให้ฝ่ายตรงข้ามตอบกลับ

ผลการแข่งขันในครั้งนี้คณะผู้พัฒนาได้รับรางวัลอันดับที่ 3 จากทั้งหมด 10 ทีมที่ส่งเข้าแข่งขัน

คำสำคัญ: บททฤษฎีประติสัมพันธ์ เกมต่อสู้

Abstract

Fighting game is a kind of computer games. There are two players in the game. To win this game, players have to attack and defense. An example of the most popular fighting game is the Street Fighter that allow players to choose character, and then having a fight with BOTs that have their own style of fighting, which had set from a game developer. The loser is a player who left all energy power. For this reason, researchers attend to the

Fighting Game AI Competition to develop more intelligence AI BOTs to be a winner of the competition.

The researchers use Java and Artificial Intelligence (AI) algorithm to create a character and fighting techniques. A continuously strike, which is not allow a rival to counter, is a key of success.

In the Fighting Game AI Competition, the researchers got the third place prize.

Keyword: AI BOTs, Fighting Game.

1. บทนำ

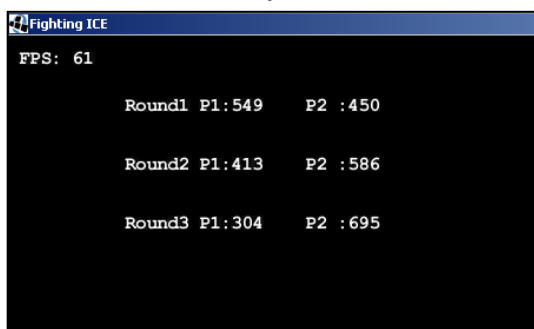
คณะผู้พัฒนาได้เข้าร่วมแข่งขัน Fighting Game AI Competition ที่จัดขึ้นในการประชุมวิชาการระดับนานาชาติ “2013 IEEE 2nd Global Conference on Consumer Electronics” ณ Intelligent Computer Entertainment Laboratory, Ritsumeikan University ประเทศญี่ปุ่น โดยมีวัตถุประสงค์เพื่อศึกษาในการพัฒนาและทดสอบอัลกอริทึมของ BOTs ผู้เข้าร่วมแข่งขัน สำหรับการแข่งขันครั้งนี้เราได้แรงผลักดันจากการทำอะไรให้ BOTs นั้นสามารถเป็นผู้ชนะในการเล่น ซึ่งมีกฎที่ใช้ในการแข่งขันดังนี้

1. กำหนดให้ใช้ตัวละครที่ให้ คือ KFM-prototype
2. การแข่งขันใน 1 ครั้งจะให้สู้กัน 3 รอบโดยแต่ละรอบ จะใช้เวลา 60 วินาที
3. ในการแข่งขันนั้น ตำแหน่งของตัวละครจะถูกกำหนดให้อยู่ในท่าเริ่มต้นทุกครั้ง
4. สูตรการคำนวณคะแนน

$$\frac{\text{Opponent Character's HP}}{\text{Player Character's HP} + \text{Opponent Character's HP}} \times 100 \quad (1)$$
5. หลังจากการแข่งขันครบ 3 รอบแล้วใครมีคะแนน HP สูงสุดจะเป็นผู้ชนะ

วัตถุประสงค์

- เพื่อพัฒนา AI BOTs ให้มีความฉลาดมากยิ่งขึ้น ที่จะได้เป็นฝ่ายชนะ โดยใช้หลักการของ AI มาช่วยในการพัฒนา
- เป้าหมายของการแข่งขันครั้งนี้ การที่จะได้มาซึ่งชัยชนะในครั้งนี้ คือ ผู้แข่งขันจำเป็นต้องมีพลังชีวิตเหลือมากที่สุด เมื่อจบการสู้ในแต่ละรอบ ดังภาพที่ 1



ภาพที่ 1: หน้าจอสรุปผลการแข่งขันในแต่ละครั้ง

- คณะผู้พัฒนาได้วิเคราะห์เพื่อหาแนวทางในการที่จะได้มาซึ่งชัยชนะ ดังนั้นเราจึงได้พัฒนา AI BOTs ด้วย Rule-based คือมีลักษณะให้มีการตอบสนองตามเงื่อนไขบางประการ

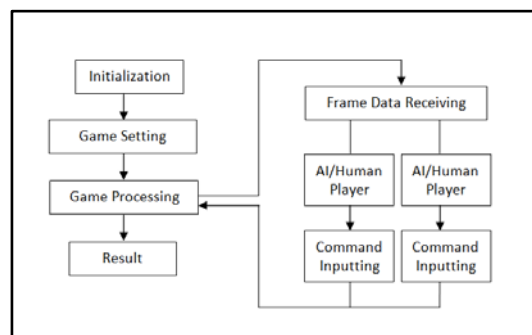
2. งานวิจัยที่เกี่ยวข้อง

ในการสร้างตัวละครเพื่อใช้ต่อสู้กับคู่ต่อสู้ในเกมต่อสู้ (Fighting Game) สิ่งสำคัญคือจะต้องสร้างกฎการต่อสู้ (Rule based) ให้กับตัวละครนั้น เพื่อให้สามารถเรียนรู้รูปแบบการต่อสู้ของคู่ต่อสู้ จากนั้นสามารถเลือกวิธีการจู่โจมและป้องกันได้อย่างเหมาะสม โดยผู้สร้างต้องเขียนอัลกอริทึมที่มีกลยุทธ์การต่อสู้เป็นแบบไดนามิก ที่มีความยืดหยุ่น สามารถเรียนรู้และตอบสนองได้ตามสถานการณ์ที่กำลังดำเนินอยู่ได้อย่างดี[1] ซึ่งสำหรับการต่อสู้แล้วต้องคำนึงถึงชัยชนะที่จะได้รับ โดยตัวละครจะต้องถูกกำหนดการต่อสู้โดยใช้ไดนามิกสคริปต์(Dynamic Script) ที่มีรูปแบบง่าย แต่มีประสิทธิภาพในการเลือกแนวทางการต่อสู้[2] เช่น เกม NEVERWINTER NIGHTS ที่นำหลักการไดนามิกสคริปต์(Dynamic Script) มาใช้ในการพัฒนา [3]

สำหรับการแข่งขันเกมต่อสู้ (Fighting Game Artificial Intelligence Competition Platform) ซึ่งการแข่งขันพัฒนาการควบคุมตัวละคร (AI Controller) โดยใช้ภาษาจาวา ซึ่งจัดโดย Intelligent Computer Entertainment Lab of Ritsumeikan University [4] นั้น มีทักษะการต่อสู้ซึ่งแบ่งเป็น 3 ส่วนคือ 1) การเริ่มต้น 2)การโจมตี และ 3) การกลับสู่สถานะปกติ โดยผู้ออกแบบอัลกอริทึมสามารถสร้างท่าโจมตีและการเคลื่อนไหวในรูปแบบเฉพาะตัวได้ [5] สำหรับแต่ละคู่ต่อสู้จะแข่งขันกัน 3 รอบ รอบละ 60 วินาที จากนั้นจะคำนวณด้วยสูตรซึ่งใช้ค่า HP ที่เหลือของคู่ต่อสู้ทั้งสองมาคำนวณ [6] ดังนั้นกลยุทธ์การแข่งขันจะต้องเน้นการโจมตีคู่ต่อสู้ และป้องกันการโจมตีให้ได้มากที่สุดภายในระยะเวลาที่กำหนด

3. ขั้นตอนการพัฒนา

1. คณะผู้พัฒนาได้ศึกษาแพลตฟอร์มของเกมต่อสู้ ที่แสดงความสัมพันธ์ระหว่าง AIs หรือ Humans ดังภาพที่ 2



ภาพที่ 2: แพลตฟอร์มของเกม

ขั้นตอนการทำงานของเกมมีดังนี้

- **Initialization** เป็นการเริ่มต้นในการเตรียมความพร้อมของอุปกรณ์คอมพิวเตอร์ และชุดคำสั่งที่ใช้ในการพัฒนา AI BOTs นั้นพร้อมที่จะเริ่มทำงาน
- **Game Setting** เป็นการติดตั้งโปรแกรมที่ใช้ในการพัฒนามีดังนี้
 - ติดตั้ง JavaSDK7
 - ติดตั้ง Eclipse Version 4.2.2

- **Game Processing** กระบวนการทำงานของเกม โดยที่ AI หรือ Human Player นั้นจะได้รับข้อมูลต่อเฟรมของเกมเพื่อที่จะควบคุมตัวละคร (Character) ที่สร้างขึ้น โดยที่ 1 เฟรม มีเวลาเท่ากับ 16.66 มิลลิวินาที ซึ่งใน 1 รอบของการแข่งขันนั้นจะมีการทำงานดังนี้

- ระบบจะทำการ Clear ข้อมูลในรอบสุดท้ายของเกม
- กำหนดค่าเริ่มต้นของ HP Energy และ Speed ของผู้เล่น ให้มีค่าเท่ากับ ศูนย์
- กำหนดตำแหน่งเริ่มต้นให้แก่ผู้เล่นทั้ง 2 ฝ่าย โดยกำหนดให้ ผู้เล่นที่1(P1) อยู่ด้านซ้าย และ ผู้เล่นที่2(P2) อยู่ด้านขวา
- กำหนดให้ผู้เล่นทั้ง 2 อยู่ในท่า(Skill) เดียวกันพร้อมที่จะต่อสู้
- มีเวลาในการแข่งขัน 60 วินาทีใน 1 รอบ
- จากนั้นระบบจะเริ่มทำงานตามคำสั่งที่ผู้พัฒนาได้กำหนดให้ AI BOTs

- **Result** คือผลคะแนน HP ที่คงเหลือในการแข่งขันแต่ละรอบ

2. คณะผู้พัฒนาได้ศึกษาและวิเคราะห์การออกท่า(Skill) ของ Character ที่จะนำมาใช้ในการสร้าง AI BOTs เพื่อใช้ในการจู่โจมและตอบรับคู่ต่อสู้ดังภาพที่ 3

Example: STAND_A frameNumber = 18 frames

Image																		
Action on/off	off	off	off	off	off	off	on	on	off	off	off	off	off	off	off	off	off	off
Frame	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

Startup = 1-6 frames Active = 7-8 frames Recovery = 9-18 frames Cancelable = 14-18 frames

ภาพที่ 3: แสดงการทำงานของท่า(Skill) ต่อคู่

ซึ่งแต่ละท่า (Skill) มีทักษะการต่อสู้ซึ่งแบ่งเป็น 3 ส่วนคือ

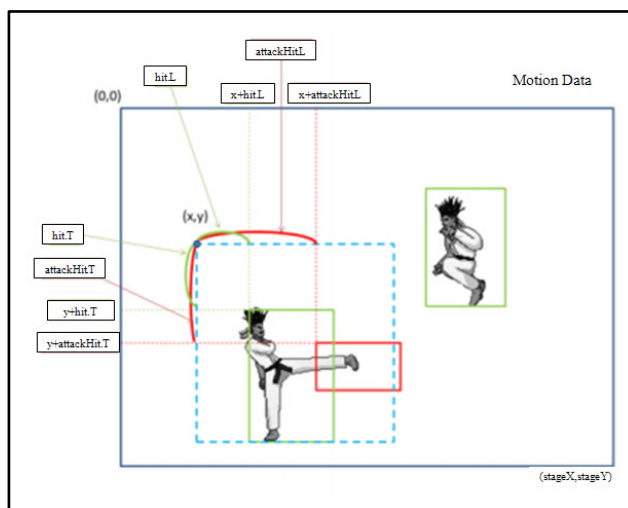
- Startup จะเริ่มเฟรมที่ 1 - 6 ใช้สำหรับเตรียมท่าก่อนเริ่มออกท่าเพื่อต่อสู้กับฝ่ายตรงข้าม
- Active เริ่มเฟรมที่ 7 - 8 เป็นช่วงสำหรับการโจมตีคู่ต่อสู้

- Recovery ใช้สำหรับการกลับสู่สถานะปกติ เริ่มใช้ตั้งแต่เฟรมที่ 9 - 18 ซึ่งบางท่า (Skill) สามารถที่จะใช้คำสั่ง Cancelable ได้

3. ศึกษาข้อมูลของ Character ซึ่งลักษณะของ Character จะมีลักษณะที่สำคัญดังนี้

- Hit Point (HP) คือคะแนนในการโจมตีของ Character ซึ่งคะแนนจะถูกตั้งค่าให้เป็นศูนย์ก่อนเริ่มการแข่งขันในรอบนั้นๆ และเมื่อถูกคู่ต่อสู้โจมตีคะแนนก็จะลดลง เมื่อการแข่งขันสิ้นสุดลงระบบจะคำนวณคะแนนคงเหลือของทั้ง 2 ฝ่ายเพื่อหาผู้ชนะในแต่ละรอบ
- Energy คือพลังที่ใช้ในแต่ละ Skill ซึ่งจะถูกตั้งค่าให้เป็นศูนย์ก่อนเริ่มการต่อสู้ในรอบนั้น พลังจะเพิ่มขึ้นก็ต่อเมื่อได้ใช้ Skill จู่โจมคู่ต่อสู้ ในบางครั้งจำเป็นต้องใช้พลังที่มากพอสำหรับ Skill บางท่า เพราะพลังอาจจะถูกนำไปใช้จนหมด ดังนั้นในการที่จะสร้าง AI BOTs ขึ้นคณะผู้พัฒนา

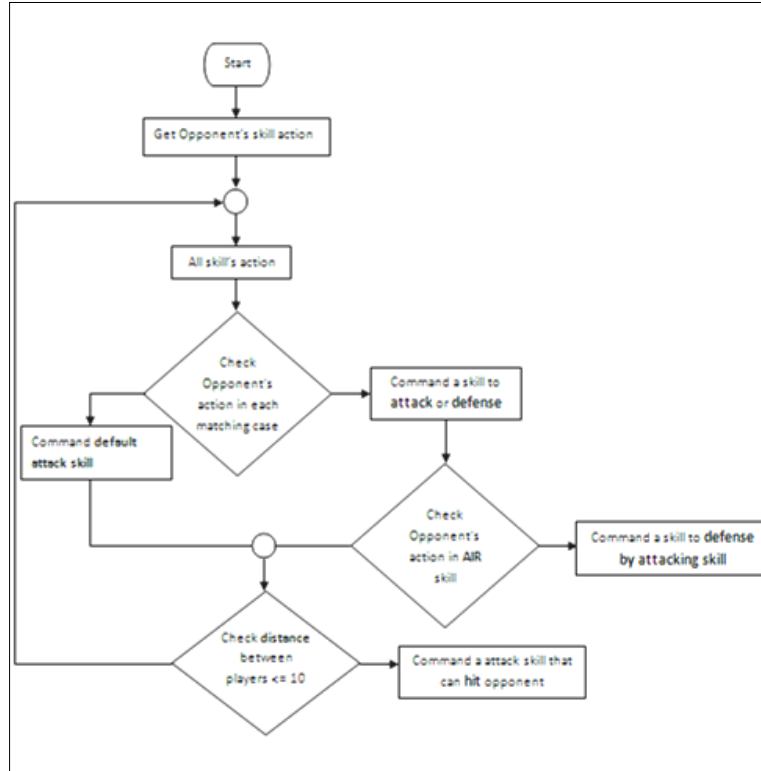
จำเป็นต้องศึกษาข้อมูลต่างๆ เช่นวิธีการเคลื่อนย้าย ระยะห่างระหว่างคู่ต่อสู้ พิกัดที่จะสามารถเข้าโจมตีได้ เพื่อให้เสียคะแนนน้อยที่สุด ดังภาพที่ 4



ภาพที่ 4: Motion Data

4. คณะผู้พัฒนาได้ทำการสร้าง AI BOTs มีทักษะในการโจมตีฝ่ายตรงข้ามอย่างต่อเนื่องโดยไม่เปิดโอกาสให้ตอบกลับ ด้วยการตรวจสอบท่าทางของฝ่ายตรงข้าม จากนั้น BOTs จะโจมตีกลับด้วยท่าทางที่จะได้

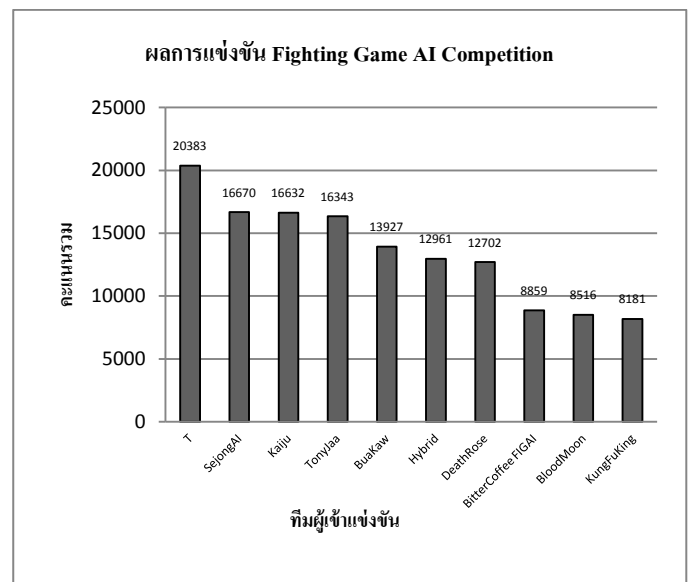
คะแนนมากกว่า หรือใช้ท่าในการป้องกัน เพื่อให้เสียคะแนนน้อยที่สุด โดยแสดงในรูปแบบผังงาน (Flowchart) การพัฒนา AI BOTs ดังภาพที่ 5



ภาพที่ 5: ผังงาน (Flowchart) การพัฒนา AI BOTs

4. ผลการวิเคราะห์

คณะผู้พัฒนาได้ทำการเก็บรวบรวมข้อมูลการนำ AI BOTs มีชื่อว่า Kaiju ที่ใช้ในการแข่งขันครั้งนี้ โดยวิเคราะห์จากผลรวมคะแนนทั้งหมดที่ได้จากการต่อสู้กับทีมอื่นๆ จำนวน 10 ทีม ผลที่ได้รับคือ ทีมของผู้พัฒนา(Kaiju)ได้คะแนนรวมเป็นอันดับที่ 3 ดังภาพที่ 6 เนื่องจากทีมที่ชนะเลิศเป็นอันดับที่ 1(T) นั้นมีการโจมตีฝ่ายตรงข้ามด้วยการใช้ทักษะการวัดตำแหน่งระยะห่างระหว่างและท่าทาง(Skill) ของคู่ต่อสู้ และ BOTs จะทำการเลือกท่าทางที่ดีที่สุดเพื่อโต้ตอบกลับ ส่วนทีมที่ 2 จะเน้นท่าทาง(Skill) การชนหรือกระแทกเพื่อทำให้คู่ต่อสู้ล้มลง(Knock down) ด้วยเหตุนี้ BOTs ของทีมผู้พัฒนาจึงไม่สามารถเอาชนะทั้ง 2 ทีมได้



ภาพที่ 6: ผลการแข่งขัน Fighting Game AI Competition

5.สรุปแนวทางในการพัฒนาต่อในอนาคต

จากการพัฒนา AI BOTs ในครั้งนี้โดยใช้วิธีการสร้างกฎการต่อสู้ (Rule based) ให้กับ BOTs โดยมีเป้าหมายในการที่จะเป็นผู้ชนะในการต่อสู้ได้ จากการวิเคราะห์ได้ข้อสรุปว่า ผู้แข่งขันจะต้องทำการป้องกันตนเองและรอจังหวะเพื่อโต้ตอบจึงจะเป็นผู้ชนะในการแข่งขัน

ผู้พัฒนาจึงได้ทำการพัฒนา AI BOTs ที่มีลักษณะการจู่โจมดังที่กล่าวมาแล้วข้างต้นและทำการทดสอบพร้อมกับเก็บรวบรวมสถิติการชนะและแพ้เมื่อผู้แข่งขันใช้ AI BOTs นี้ และได้ผลเป็นที่น่าพอใจ ซึ่งในการแข่งขันครั้งนี้ คณะผู้พัฒนาได้รับรางวัลอันดับที่ 3 จากทั้งหมด 10 ทีมที่ส่งเข้าแข่งขัน

ในการพัฒนา AI BOTs สำหรับการเข้าแข่งขันในครั้งต่อไป คณะผู้พัฒนาจะพัฒนา AI BOTs ให้รองรับการแข่งขันที่มีรูปแบบของคู่ต่อสู้ที่หลากหลายมากขึ้น ด้วยการนำหลักเนื่องจากคู่ต่อสู้มีการจู่โจมด้วยทักษะที่หลากหลาย ดังนั้นคณะผู้พัฒนาจะต้องออกแบบ AI BOTs ให้สามารถทั้งจู่โจมและป้องกันตัว จากฝ่ายตรงข้ามได้เป็นอย่างดี เพื่อให้เป็นฝ่ายชนะ

กิตติกรรมประกาศ

ขอกราบขอบพระคุณดร. กิ่งกาญจน์ สุขคณาภิบาล และดร. วรวัฒน์ เชิญสวัสดิ์ อาจารย์ที่ปรึกษาในการส่งผลงานเข้าแข่งขันในครั้งนี้ และคอยให้คำปรึกษา ให้แนวคิด และช่วยตรวจสอบข้อบกพร่องพร้อมแนะแนวทางแก้ไข

เอกสารอ้างอิง

- [1] Ricciardi, Antonio, and Patrick Thill. "Adaptive AI for Fighting Games." December 12.200 (2008): 8.
- [2] Ruck Thawonmas and Syota Osaka, "A Method for Online Adaptation of Computer-Game AI Rulebase," ACM SIGCHI International Conference on Advances in Entertainment Technology 2006, Hollywood, USA, Jun, 2006.
- [3] Peter Spronck, Marc Ponsen, Ida Sprinkuizen-Kuyper and Eric Postma "Adaptive Game AI with Dynamic Scripting" 2005 Kluwer Academic Publisher
- [4] Fighting Game Artificial Intelligence Competition <http://www.ice.ci.ritsumeai.ac.jp/~ftgaic/index.htm>

- [5] Feiyu Lu, Kaito Yamamoto, Luis H. Nomura, Syunsuke Mizuno, YoungMin Lee, and Ruck Thawonmas, "Fighting Game Artificial Intelligence Competition Platform," *Proceeding of the 2nd IEEE Global Conference on Consumer Electronics (GCCE 2013)*, Tokyo, Japan, pp. 320-323, Oct. 1-4, 2013.
- [6] The Rule of the Game, <http://www.ice.ci.ritsumeai.ac.jp/~ftgaic/index-3.html>

การเปรียบเทียบความน่าเชื่อถือในการวัดมุมความโค้งของกระดูกสันหลัง

A Comparison of Reliability in Measuring Spinal Curvature

อภิชนา เกียรติอุบลไพบุณย์ (Apichaya Kiartubonpaiboon)¹ และ วรรณัญ วงษ์เสรี (Waranyu Wongseree)²

สาขาวิชาเทคโนโลยีการจัดการระบบสารสนเทศ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยมหิดล

Emails: ¹apic5@hotmail.com, ²waranyu1979@hotmail.com

บทคัดย่อ

การศึกษาค้นคว้าครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบค่าความน่าเชื่อถือของวิธีการวัดมุมความโค้งกระดูกสันหลังสามวิธี คือ วิธีการวัดแบบ Ferguson, Cobb และ พหุนาม ซึ่งวิธีการวัดแบบ Ferguson เป็นวิธีการวัดที่เก่าแก่ที่สุด ส่วนวิธีการวัดแบบ Cobb เป็นวิธีการวัดที่มีความนิยมสูงสุด และ วิธีการวัดแบบ พหุนาม เป็นหนึ่งในแบบจำลองทางคณิตศาสตร์แรกที่ถูกเสนอในการวัดมุมความโค้งกระดูกสันหลัง โดยศึกษาวิธีการวัดมุมความโค้งกระดูกสันหลังด้านหน้า (AP-View) จากภาพฉายรังสี 2 มิติ ในผู้ป่วยโรคกระดูกสันหลังโรงพยาบาลเลิดสิน จำนวน 30 ภาพ วัดโดยนักวิจัยจำนวน 3 คน ที่ผ่านการอบรมการกำหนดจุดขอบกระดูกสันหลัง (Landmarks) จากแพทย์ผู้เชี่ยวชาญ สถิติที่ใช้คือ Intra class Correlation Coefficient (ICC) เป็นการเปรียบเทียบค่าความน่าเชื่อถือของผู้ประเมิน (Intra-rater reliability) โดยใช้แบบจำลอง ICC (3,1) และ ความน่าเชื่อถือระหว่างผู้ประเมิน (Inter-rater reliability) โดยใช้แบบจำลอง ICC (2,1) จากผลการศึกษาพบว่า ค่า Intra-rater reliability ของวิธีการวัดแบบ Ferguson, Cobb และ พหุนาม มีค่าเท่ากับ 0.968, 0.950 และ 0.910 ตามลำดับ ส่วนค่า Inter-rater reliability ของวิธีการวัดแบบ Ferguson, Cobb และ พหุนาม มีค่าเท่ากับ 0.477, 0.659 และ 0.407 ตามลำดับ ผลการศึกษานี้บ่งชี้ว่า วิธีการวัดมุมความโค้งกระดูกสันหลังที่มีค่าความน่าเชื่อถือของผู้ประเมินมากที่สุดคือ Ferguson และ วิธีการวัดมุมความโค้งกระดูกสันหลังที่มีค่าความน่าเชื่อถือระหว่างผู้ประเมินมากที่สุดคือ Cobb

คำสำคัญ: วิธีการวัดแบบ Ferguson; วิธีการวัดแบบ Cobb; วิธีการวัดแบบพหุนาม; ความน่าเชื่อถือในผู้ประเมิน (ICC); ภาพกระดูกสันหลังด้านหน้า (AP-View);

การกำหนดจุด; ค่าความน่าเชื่อถือตัวผู้ประเมิน; ค่าความน่าเชื่อถือระหว่างผู้ประเมิน

Abstract

The purposes of this study were to assess intra-rater and inter-rater reliability of 3 methods for measuring spinal curvature in 30 anteroposterior radiograph 2D view (AP view) with spinal patients of Lerdsin hospital during the year from 2004 to 2011. Cobb, Ferguson and Polynomial method were compared. Ferguson method is the traditional method and Cobb method is probably the most popular, while polynomial method is one of the first documented mathematical models for sagittal spinal curvature. Intraclass correlation coefficient (ICC) was used to calculate the inter-rater reliability (ICC model 2, 1) and intra-rater reliability (ICC model 3, 1). The statistical analysis showed that the intra-rater reliability of Ferguson, Cobb and Polynomial methods is 0.968, 0.950 and 0.910 respectively. In similar with their inter-rater reliability is 0.477, 0.659 and 0.407 respectively. The results indicated that the most reliable measurement of spinal curvature is Ferguson method, when used to assessor. And the most reliable measurement of spinal curvature is Cobb method, when using between examiners.

Keywords: Ferguson Method, Cobb Method, Polynomial Method, Intra class Correlation Coefficient (ICC), Anteroposterior radiographic view (AP-View), Landmarks, Intra-rater reliability, Inter-rater reliability

1. บทนำ

กระดูกสันหลังเป็นโครงสร้างที่แข็งแรง มีหน้าที่ปกป้องแกนไขสันหลัง ของกระดูกสันหลัง ทำหน้าที่เป็นจุดเกาะของกล้ามเนื้อของหลัง และยังเชื่อมต่อกับกระดูกสำคัญอื่นๆของร่างกาย ซึ่งอาการที่พบเห็นได้บ่อยในผู้ป่วยโรคที่เกี่ยวกับกระดูกสันหลังคือ กระดูกสันหลังคด (Scoliosis) และ อาการปวดกระดูกสันหลังเนื่องจากการเสื่อมของอายุ (Spondylosis) โดยอาการกระดูกสันหลังคด, แอ่น เป็นภาวะที่กระดูกสันหลังมีความโค้งในแนวซ้ายขวาที่ผิดปกติ นอกจากนี้อาจมีการบิดหรือหมุนออกไปจากแนวเดิมของกระดูกสันหลัง อาจเกิดจากความผิดปกติแต่กำเนิด (Congenital disorders) หรือ ระหว่างการเจริญเติบโตจากการเสื่อมของกระดูกสันหลัง รวมไปถึงหลังโก่ง (Kyphosis) ทำให้แรงกดบนหมอนรองกระดูกไม่สม่ำเสมอ เอ็นกล้ามเนื้อรับน้ำหนักตลอดเวลา ทำให้มีการเสื่อมและทำลายเนื้อเยื่อต่าง ๆ ส่งผลให้เกิดอาการปวดหลัง [1], [2] ส่วนอาการปวดกระดูกสันหลังเนื่องจากการเสื่อมของอายุ เป็นภาวะที่เกิดจากอายุที่มากขึ้น โดยในปัจจุบันโรคกระดูกสันหลังคด และ โรคกระดูกสันหลังเคลื่อน มีอัตราการเกิดมากขึ้น ทั้งเพศหญิง เพศชาย เด็ก ผู้ใหญ่ และ ผู้สูงอายุ มีอาการที่หนักเบาแตกต่างกัน ปัจจัยที่ก่อให้เกิดการปวดหลัง และ การเสื่อมของข้อกระดูก คือ เพศ อายุ ดัชนีมวลกาย (BMI) จะพบมากในผู้ป่วยสูงอายุ และ ผู้ป่วยที่มีดัชนีมวลกายมาก [3] มีสาเหตุมาจากการเพิ่มขึ้น หรือ ลดลงจากความโค้งของกระดูกสันหลัง เนื่องจากความโค้งของกระดูกสันหลังทำหน้าที่รับน้ำหนักของร่างกาย หากมีการเพิ่มขึ้น หรือ ลดลงจากความโค้งของกระดูกสันหลัง ส่งผลให้มีแรงกระทำต่อโครงสร้างบริเวณกระดูกสันหลังมากขึ้น ทำให้เกิดการกดทับหรือการเบียดเสียดต่อโครงสร้างบริเวณกระดูกสันหลัง เป็นสาเหตุทำให้เกิดอาการปวดหลัง [1] ซึ่งการรักษาจะต้องใช้วิธีการเอ็กซเรย์ให้ได้ภาพถ่ายรังสีกระดูกสันหลังของผู้ป่วย และนำมาประกอบการวัดมุมความโค้งของกระดูกสันหลังจากรูปภาพฉายรังสี เพื่อคำนวณมุมที่จะนำไปสู่การรักษาที่เหมาะสม และ ถูกต้อง ซึ่งงานวิจัยนี้เปรียบเทียบความน่าเชื่อถือของวิธีการวัดมุมความโค้งกระดูกสันหลังที่มีค่าความน่าเชื่อถือที่สุด ระหว่างวิธีการวัดแบบ Ferguson, Cobb และ พหุนาม

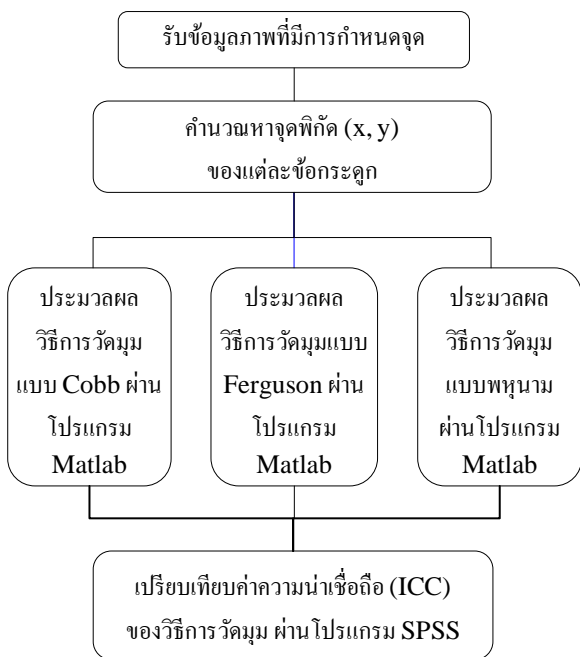
2. งานวิจัยที่เกี่ยวข้อง

การประเมินความโค้งของกระดูกสันหลังในปัจจุบันสามารถแบ่งได้เป็น 2 วิธี คือ วิธีที่มีการรुक้า (invasive techniques) และ วิธีที่ไม่มีการรुक้า (non-invasive techniques) โดยการวัดแบบที่ไม่มีการรुक้าสามารถทำได้จากการใช้อุปกรณ์ต่างๆ เช่น การวัดที่ใช้ Flexicurve ส่วนการวัดแบบรुक้า ได้แก่ การใช้ภาพถ่ายรังสี เช่น วิธีการวัดแบบ Cobb ซึ่งเป็นวิธีมาตรฐานที่ให้ผลการวัดที่มีความถูกต้องสูง [4], [5] การวัดมุมความโค้งของกระดูกสันหลังจากภาพถ่ายรังสี มีการวัดจากภาพถ่ายรังสีของกระดูกสันหลังด้านหน้า (AP-View) และ ด้านข้าง (Lateral View) ต่างมีวิธีการวัดมุมความโค้งที่หลากหลาย ซึ่งการวัดมุมในแต่ละวิธีจะใช้การวัดค่าจากขอบกระดูกฝั่งด้านหน้า (Anterior), และ ขอบกระดูกฝั่งด้านหลัง (Posterior) ด้วยการกำหนดจุด (Landmarks) มีการใช้การกำหนดจุดแบบ 4, 6, 9 และ 36 จุด โดยการกำหนดจุดแบบ 6, 9 และ 36 จุด ใช้สำหรับภาพถ่ายรังสี 2 มิติ กระดูกสันหลังด้านข้าง การกำหนดจุดแบบ 6 จุด ใช้สำหรับข้อกระดูกที่ไม่มีกระดูกงอก (Osteophytes) การกำหนดจุดแบบ 9 จุด ใช้สำหรับข้อกระดูกที่มีกระดูกงอก [6] การกำหนดจุดแบบ 36 จุด ใช้สำหรับการกำหนดรูปร่างของข้อกระดูกสันหลังเพิ่มความละเอียดรูปร่างจากการกำหนดจุดแบบ 9 จุด [7] ส่วนการกำหนดจุดแบบ 4 จุดใช้สำหรับภาพถ่ายรังสีกระดูกสันหลังด้านหน้า และ ด้านข้าง เพื่อการหาค่าความโค้งของกระดูกสันหลัง ที่มุมของขอบกระดูกทั้ง 4 จุด การกำหนดจุดจะมีค่าความคลาดเคลื่อนจากการใช้คนในการกำหนดจุดไม่สามารถกำหนดจุดค่าออกมาเหมือนเดิมทุกครั้ง ขณะที่ผู้วัดด้วยบุคคลคนเดียวมีความคลาดเคลื่อนพอสมควร หากมีจำนวนผู้วัดมากขึ้นทำให้ค่าความคลาดเคลื่อนนั้นเพิ่มตามไปด้วย

เทคนิคการวิเคราะห์ทางสถิติที่ใช้ในการเปรียบเทียบความน่าเชื่อถือคือ Intra class Correlation Coefficient (ICC) มีการศึกษาวิจัยเปรียบเทียบความน่าเชื่อถือวิธีการวัดมุมความโค้งกระดูกสันหลังจากภาพถ่ายรังสี 2 มิติ กระดูกสันหลังด้านหน้า โดยเปรียบเทียบค่าความน่าเชื่อถือของผู้ประเมิน และ ระหว่างผู้ประเมิน พบว่าวิธีการวัดแบบ Cobb มีความน่าเชื่อถือมากที่สุด [8] จึงเป็นที่มาของการเปรียบเทียบความน่าเชื่อถือของวิธีการวัดมุมความโค้งกระดูกสันหลังของคนไทย

3. ขั้นตอนวิธีการดำเนินการวิจัย

3.1 ขั้นตอนการดำเนินการ



ภาพที่ 1 : ขั้นตอนการดำเนินการวิจัย

ขั้นตอนการดำเนินการวิจัย รับข้อมูลภาพฉายรังสี 2 มิติ กระดูกสันหลังด้านหน้าที่มีการกำหนดจุดโดยผู้ประเมิน เพื่อทำการคำนวณจุดพิกัด (x, y) ของกระดูกสันหลังแต่ละข้อกระดูก L1 – L5 แล้วนำจุดพิกัด (x, y) ดังกล่าวมาประมวลผลด้วยวิธีการวัดมุมความโค้งกระดูกสันหลังแบบ Cobb, Ferguson และ พหุนาม ผ่านโปรแกรม Matlab ค่าที่ได้ คือ มุมองศาความโค้งกระดูกสันหลัง จากนั้นเปรียบเทียบค่าความน่าเชื่อถือ Intra class Correlation Coefficient (ICC) ของวิธีการวัดมุมความโค้งทั้ง 3 วิธี ผ่านโปรแกรม SPSS ค่าที่ได้ คือ ค่าความน่าเชื่อถือของผู้ประเมิน (Intra-rater reliability) และ ค่าความน่าเชื่อถือระหว่างผู้ประเมิน (Inter-rater reliability) ของวิธีการวัดมุมความโค้งกระดูกสันหลังแบบ Cobb, Ferguson และ พหุนาม ดังภาพที่ 1

3.2 ศึกษาและรวบรวมข้อมูล

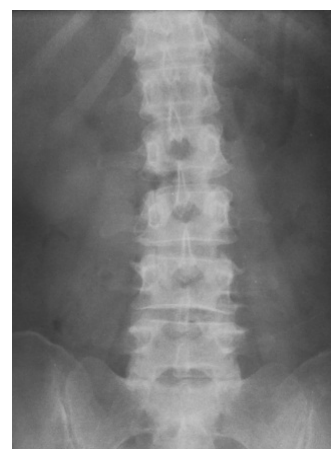
จากการศึกษาและรวบรวมข้อมูลภาพฉายรังสี 2 มิติ กระดูกสันหลังด้านหน้า แบ่งเป็น ภาพที่ไม่มีความชัดเจนของขอบกระดูกสันหลังส่วนเอว กระดูกชั้นที่ L1 - L5 ดังภาพที่ 2 อาจก่อให้เกิดข้อผิดพลาดในการวัดจากการกำหนดจุด และ ภาพที่มีความชัดเจนของขอบกระดูกสันหลังส่วนเอว กระดูกชั้นที่ L1 -

L5 ดังภาพที่ 3 เพื่อสะดวกต่อผู้วัดในการกำหนดจุด ทำให้ลดข้อผิดพลาดในการวัดมุมความโค้งกระดูกสันหลังได้

กลุ่มตัวอย่างการศึกษาวิจัยครั้งนี้เป็นนักวิจัยที่ได้รับการอบรมการกำหนดจุดขอบกระดูกสันหลังจากแพทย์ผู้เชี่ยวชาญจำนวน 3 คน และ ข้อมูลภาพฉายรังสี 2 มิติ กระดูกสันหลังด้านหน้า ในผู้ป่วยโรคกระดูกสันหลัง โรงพยาบาลเลิศสิน จำนวน 30 ภาพ โดยใช้ภาพที่มีความชัดเจนของขอบกระดูกสันหลังส่วนเอว กระดูกชั้นที่ L1 - L5 ในการกำหนดจุดกระดูกสันหลังจากขอบของกระดูกสันหลังส่วนเอว เพื่อใช้ในการวัดมุมความโค้งกระดูกสันหลังด้วยวิธีแบบ Ferguson, Cobb และ พหุนาม แล้วจึงนำมาวิเคราะห์ทางสถิติเพื่อหาความน่าเชื่อถือของวิธีการวัดมุมความโค้งกระดูกสันหลัง



ภาพที่ 2 : ภาพฉายรังสี 2 มิติ กระดูกสันหลังด้านหน้า (AP-View) ซึ่งไม่มีความชัดเจนของขอบกระดูกสันหลังส่วนเอวข้อที่ L1-L5



ภาพที่ 3 : ภาพฉายรังสี 2 มิติ กระดูกสันหลังด้านหน้า (AP-View) ซึ่งมีความชัดเจนของขอบกระดูกสันหลังส่วนเอวข้อที่ L1-L5

3.3 การกำหนดจุดกระดูกสันหลัง

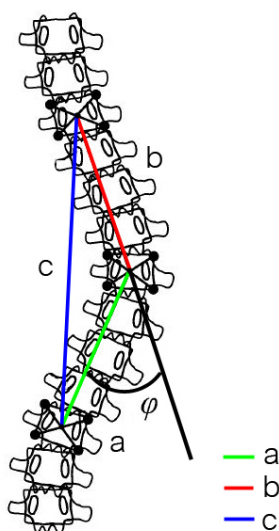
ภาพถ่ายรังสี 2 มิติ กระดูกสันหลังด้านหน้า แสดงให้เห็นถึงรูปทรงขอบ และ มุมทั้ง 4 มุม ของข้อกระดูกสันหลังส่วนเอว กระดูกชั้นที่ L1 – L5 โดยทำการกำหนดจุดในกระดูกแต่ละชั้น จำนวน 4 จุด มีช่วงข้อกระดูกฝั่งด้านหลัง คือ จุดที่ 1 และ จุดที่ 2 ส่วนช่วงข้อกระดูกฝั่งด้านหน้า คือ จุดที่ 3 และ จุดที่ 4 ดังภาพที่ 4



ภาพที่ 4 : ภาพกำหนดจุดกระดูกสันหลัง

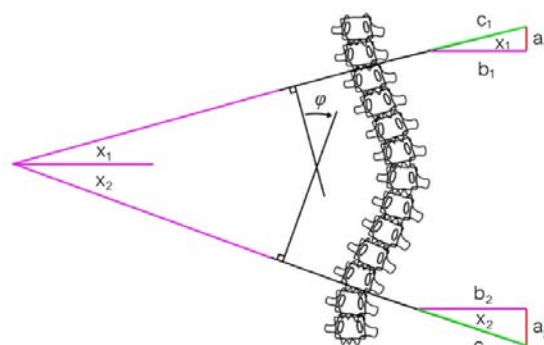
3.4 วิธีการวัดมุมความโค้งของกระดูกสันหลัง

การวัดมุมแบบ Ferguson จะมีการกำหนดจุดทั้งหมด 12 จุด โดยแบ่งจุดตามขอบของข้อกระดูกสันหลังส่วนเอว เป็น ข้อบน, ข้อตรงกลาง และ ข้อล่าง กำหนดจุดข้อละ 4 จุด ดังภาพที่ 5 จากนั้นหาจุดศูนย์กลางของกระดูกแต่ละข้อแล้วลากเส้นเชื่อมต่อกระดูกทั้งสามข้อ และ ทำการวัดความโค้งจากมุมที่เกิดจากการตัดกันของด้าน a และ b จะได้มุมความโค้งกระดูกสันหลังเป็นค่าของมุมแอลฟา (ϕ)



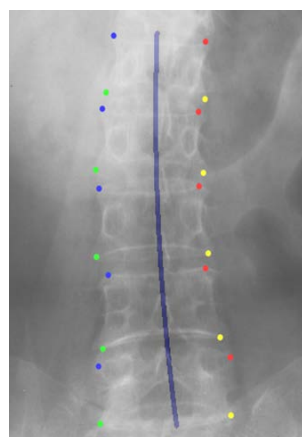
ภาพที่ 5 : วิธีการวัดแบบ Ferguson [8], [9]

การวัดมุมแบบ Cobb จะมีการกำหนดจุดทั้งหมด 4 จุด โดยโดยแบ่งจุดตามขอบของข้อกระดูกสันหลังส่วนเอว L1 – L5 เป็น 2 จุดที่ขอบบนของกระดูกชั้นบนสุด และ 2 จุดที่ขอบล่างของกระดูกชั้นล่างสุด ดังภาพที่ 6 จะได้ความยาวด้าน a, b และ c จากนั้นทำการวัดมุมความโค้งที่เกิดจากการลากเส้นตัดกันของกระดูกชั้นบนและกระดูกชั้นล่าง โดยใช้สมบัติของสามเหลี่ยมคล้าย และ กฎของไซน์ จากนั้นใช้สมบัติของมุมบนเส้นตรง จะได้มุมความโค้งกระดูกสันหลังเป็นค่าของมุมแอลฟา (ϕ)



ภาพที่ 6 : วิธีการวัดแบบ Cobb [8]

การวัดมุมพหุนามจะมีการกำหนดจุดทั้งหมด 20 จุด โดยแบ่งจุดตามขอบของข้อกระดูกสันหลังส่วนเอว L1 – L5 แต่ละข้อมี 4 จุด จากนั้นนำพิสัยของกระดูกทุกข้อมาสร้างสมการพหุนาม ด้วยวิธีการวิเคราะห์การถดถอย ในการศึกษาใช้พหุนามอันดับสองในการสร้างสมการ จากนั้นนำสมการมาคำนวณหาค่ามุมระหว่างกระดูกชั้นบนและชั้นล่าง ดังภาพที่ 7



ภาพที่ 7 : วิธีการวัดแบบพหุนาม

4. ผลการดำเนินการวิจัย

ผลการเปรียบเทียบค่าความน่าเชื่อถือของวิธีการวัดมุมความโค้งของกระดูกสันหลังด้านหลังด้านหน้า (AP-View) จากภาพฉายรังสี 2 มิติ ในผู้ป่วยโรคกระดูกสันหลัง โรงพยาบาลเลิดสิน จำนวน 30 ภาพ ของผู้ประเมิน ICC (3, 1) และ ระหว่างผู้ประเมิน ICC (2, 1) ในรูปแบบข้อมูลที่แน่นอน โดยผู้ประเมินทำการวัดมุมความโค้ง จำนวน 3 ครั้งต่อภาพ และ ระหว่างผู้ประเมินทำการวัดมุมความโค้ง จำนวน 3 คน คนละ 1 ครั้งต่อภาพ ในตารางที่ 1 แสดงให้เห็นว่า ค่าความน่าเชื่อถือของผู้ประเมิน วิธีแบบ Ferguson มีค่าความเชื่อถือสูงสุด ICC (3,1) = 0.969 และ วิธีวัดมุมแบบพหุนามมีค่าความน่าเชื่อถือต่ำสุด ICC (3,1) = 0.912 ส่วนค่าความน่าเชื่อถือระหว่างผู้ประเมินวิธีแบบ Cobb มีค่าความน่าเชื่อถือสูงสุด ICC (2,1) = 0.644 และ วิธีวัดมุมพหุนามมีค่าความน่าเชื่อถือต่ำสุด ICC (2,1) = 0.392

ผลการเปรียบเทียบค่าความน่าเชื่อถือของวิธีการวัดมุมความโค้งของกระดูกสันหลังของผู้ประเมิน ICC (3, 1) และ ระหว่างผู้ประเมิน ICC (2, 1) รูปแบบข้อมูลสอดคล้องดังตารางที่ 2 แสดงให้เห็นว่า ค่าความน่าเชื่อถือของผู้ประเมิน วิธีแบบ Ferguson มีค่าความเชื่อถือสูงสุด ICC (3,1) = 0.968 และ วิธีวัดมุมแบบพหุนามมีค่าความน่าเชื่อถือต่ำสุด ICC (3,1) = 0.910 ส่วนค่าความน่าเชื่อถือระหว่างผู้ประเมินวิธีแบบ Cobb มีค่าความน่าเชื่อถือสูงสุด ICC (2,1) = 0.659 และ วิธีวัดมุมพหุนามมีค่าความน่าเชื่อถือต่ำสุด ICC (2,1) = 0.407

ผลจากตารางที่ 1 และ 2 สรุปได้ว่าวิธีการวัดมุมความโค้งกระดูกสันหลังแบบ Ferguson มีความน่าเชื่อถือของผู้ประเมินมากกว่า วิธีแบบ Cobb และ วิธีแบบพหุนาม ส่วนวิธีการวัดมุมความโค้งกระดูกสันหลังแบบ Cobb มีความน่าเชื่อถือระหว่างผู้ประเมินมากกว่า วิธีแบบ Ferguson และ วิธีแบบพหุนาม

ตารางที่ 1: การเปรียบเทียบค่าความเชื่อถือของวิธีการวัดมุมความโค้งกระดูกสันหลังของผู้ประเมิน ICC (3, 1) และ ระหว่างผู้ประเมิน ICC (2, 1) รูปแบบข้อมูลที่แน่นอน

วิธีการวัดมุม	ความแน่นอน (Absolute agreement)		
	Ferguson	Cobb	พหุนาม
ICC (3, 1) Intra	0.969	0.952	0.912
ICC (2, 1) Inter	0.481	0.644	0.392

ตารางที่ 2: การเปรียบเทียบค่าความเชื่อมั่นของวิธีการวัดมุมความโค้งกระดูกสันหลังของผู้ประเมิน ICC (3, 1) และ ระหว่างผู้ประเมิน ICC (2, 1) รูปแบบที่ข้อมูลสอดคล้อง

วิธีการวัดมุม	ความสอดคล้อง (Consistency)		
	Ferguson	Cobb	พหุนาม
ICC (3, 1) Intra	0.968	0.950	0.910
ICC (2, 1) Inter	0.477	0.659	0.407

5. สรุป

การศึกษาวิจัยครั้งนี้พบว่า ค่าความน่าเชื่อถือที่แน่นอนและสอดคล้องกัน มีค่าใกล้เคียงกัน แสดงให้เห็นว่าวิธีการวัดมุมความโค้งกระดูกสันหลังแบบ Cobb และ Ferguson มีความน่าเชื่อถือมากกว่าวิธีการวัดมุมความโค้งกระดูกสันหลังแบบพหุนาม ซึ่งวิธีการวัดมุมความโค้งกระดูกสันหลังแบบพหุนามมีความเป็น Global มากกว่า ซึ่งเป็นการวัดที่ประมวลผลของข้อมูลพิกัดจุดจากการกำหนดจุดของข้อกระดูกสันหลังส่วนเอวกระดูกชั้นที่ L1 - L5 ทั้งหมด จำนวน 20 จุด แตกต่างจากวิธีการวัดมุมความโค้งกระดูกสันหลังแบบ Cobb และ วิธีการวัดมุมความโค้งกระดูกสันหลังแบบ Ferguson ซึ่งมีลักษณะเป็น Local ซึ่งเป็นการประมวลผลของข้อมูลพิกัดจุดจากการกำหนดจุดของข้อกระดูกสันหลังส่วนเอวเฉพาะส่วน คือวิธีการวัดมุมความโค้งกระดูกสันหลังแบบ Ferguson ประมวลผลเฉพาะข้อกระดูกสันหลังส่วนเอวข้อบนสุด (L1) ข้อตรงกลาง (L3) และ ข้อล่างสุด (L5) ส่วนวิธีการวัดมุมความโค้งกระดูกสันหลังแบบ Cobb ประมวลผลเฉพาะข้อกระดูกสันหลังส่วนเอวข้อบนสุด (L1) และ ข้อล่างสุด (L5) ดังนั้นการวัดมุมความโค้งกระดูกสันหลังลักษณะเป็น Global จึงทำให้วัดมุมความโค้งได้น้อยกว่าลักษณะเป็น Local แต่ข้อดีของลักษณะเป็น Global คือสามารถใช้งานแก่ประเภทของความโค้งของกระดูกสันหลังได้จากการพิจารณาสัมประสิทธิ์ของสมการ และ เพื่อปรับปรุงวิธีการวัดมุมแบบพหุนามให้ดีขึ้นอาจจะปรับวิธีเป็นพหุนามแบบ Local หรือ Spline อาจจะให้ผลใกล้เคียงวิธีการวัดแบบ Cobb และ Ferguson มากขึ้น

6. กิตติกรรมประกาศ

ขอขอบคุณ นายแพทย์รัชวัช ประสาทฤทธา รองผู้อำนวยการสถาบันออร์โธปิดิกส์ แพทย์ผู้ชำนาญการทางด้านศัลยกรรมกระดูกและข้อ โรงพยาบาลเลิดสิน สำหรับการอำนวยความสะดวกข้อมูลผู้ป่วย รวมถึงข้อมูลทางการแพทย์เชิงการวัดมุมความโค้งของกระดูกสันหลัง และการกำหนดจุดขอบกระดูกสันหลังในแต่ละข้อกระดูก

เอกสารอ้างอิง

- [1] S. Esses, Spinal Deformity, Textbook of Spinal Disorders, Philadelphia: J.B. Lippincott Company, 1995.
- [2] K. Bridwell and R. DeWald, Posterior and Posterolateral Approaches to Spine, The Textbook of Spinal Surgery, 2nd ed Philadelphia: Lippincott-Raven: 193, 1997.
- [3] P. Pouletaut, H. Dalqamoni, F. Marin, and M.-C. Ho Ba Thoa, "Influence of age, gender and weight on spinal osteoarthritis in the elderly: An analysis of morphometric changes using X-ray images," IRBM Journal, vol. 31, pp. 141–147, 2010.
- [4] R. PuntumetaKul, P. Hiruntrakul, W. Premchaisawat, M. Puntumetakul, and Y. Thavornpitak, "The measurement of lumbar spinal curvature in normal Thai population aged 20-69 years using flexible ruler," J Med Tech Phy Ther, vol. 24, no.3, pp. 309-317, Sep. 2012.
- [5] T. Thaweewannakij, S. Wongsu, W. Kamruecha, J. Khaengkhan, J. Wongkuanklom, C. Konkamtan, and S. Amatachaya, "Validity and discriminative ability on physical impairment relating to kyphosis using 1.7-cm block," KGU Res. Journal, vol. 17, no. 4, pp. 660-670, 2012.
- [6] M. Gstoettner, K. Sekyra, N. Walochnik, P. Winter, R. Wachter, and C. M. Bach, "Inter and intra-observer reliability assessment of the cobb angle: manual versus digital measurement tools," Eur Spine Journal, vol. 16, no. 10, pp. 1587–1592, 2007.
- [7] D. Lee, S. Antani, Y. Chang, K. Gledhill, L. Rodney Long, and P. Christensen, "CBIR of spine X-ray images on inter-vertebral disc space and shapeprofiles using feature ranking and voting consensus," Data & Knowledge Engineering, vol. 68, pp. 1359–1369, 2009.
- [8] T. Vrtovec, F. Pernus, and B. Likar, "A review of methods for quantitative evaluation of spinal curvature," Eur Spine Journal, vol. 18, pp. 593–607, 2009.
- [9] A. C. Kittleston and L. W. LIM, "Measurement of scoliosis," American Journal of Roentgenology, vol. 108, pp. 775-777, 1970.
- [10] A. Ferguson, "The study and treatment of scoliosis," South Med Journal, vol. 23, pp. 116–120, 1930.

โปรแกรมประยุกต์บนโทรศัพท์เคลื่อนที่เพื่อนำชมเรื่องรามเกียรติ์บนกำแพงวัดพระแก้ว

Mobile device augmented reality tour application for Ramayana story on Wat Prakaew wall

กรกช เพ็งสมบูรณ์(Goragod Pangsomboon)¹ วัชรพัฐุ เลานวัฒนา(Watcharapat Laonawatthana)¹

ประวิทย์ อ่องจรรยา(Prawit Oungcharoon)¹ สภา จรรยาชัชवाल(Sapa Chanyachatchawan)²

รัฐศิลป์ รานอกphanuwat(Ratthasilp Ranokphanuwat)¹ และ เนื่องวงศ์ ทวยเจริญ(Nuengwong Tuaycharoen)¹

¹ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยธุรกิจบัณฑิตย์

²ห้องปฏิบัติการวิจัยสารสนเทศสื่อประสม ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ

p.goragod@gmail.com, {plengslowtoad, prawit_10}@hotmail.com, sapa.cha@nectec.or.th, nuengwong.tun@dpu.ac.th

บทคัดย่อ

ภาพจิตรกรรมฝาผนังเรื่องรามเกียรติ์ บนกำแพงวัดพระศรีรัตนศาสดาราม (วัดพระแก้ว) เป็นสมบัติของชาติที่เยาวชนไทยควรภาคภูมิใจ บทความนี้กล่าวถึงการพัฒนาโปรแกรมประยุกต์บนโทรศัพท์เคลื่อนที่เพื่อนำชมเรื่องรามเกียรติ์บนกำแพงวัดพระแก้ว ซึ่งส่งเสริมการเผยแพร่เรื่องราวบนภาพจิตรกรรมเรื่องรามเกียรติ์บนฝาผนังวัดพระแก้ว โดยผู้ใช้งานโทรศัพท์หรือแท็บเล็ตที่เชื่อมต่ออินเทอร์เน็ตมาถ่ายภาพบนฝาผนัง หากภาพตรงกับภาพในฐานข้อมูล เครื่องแม่ข่ายจะส่งวีดิโอรามเกียรติ์ตอนที่เกี่ยวข้องมาแสดงบนอุปกรณ์เคลื่อนที่ โดยใช้เทคนิคการรู้จำภาพ (image recognition) และการเสริมความจริง (Augmented Reality) ในการนำเสนอผลงานสื่อผสมให้มีความน่าสนใจมากขึ้น ผลการทดสอบระบบพบว่า การปรับให้โปรแกรมสกัด descriptor จากภาพในฐานข้อมูลไว้ล่วงหน้า ทำให้โปรแกรมทำงานได้เร็วขึ้น 6 เท่า และความแม่นยำในการค้นหาภาพและเวลาที่ใช้รวมทั้งระบบขึ้นอยู่กับคุณสมบัติของกล้องในอุปกรณ์เคลื่อนที่

คำสำคัญ: โปรแกรมประยุกต์บนโทรศัพท์เคลื่อนที่ วัดพระแก้ว การท่องเที่ยว การรู้จำภาพ การเสริมความจริง

Abstract

Ramayana mural on Wat Phra Kaew's wall was painted during the reign of King Rama I. Thai youths should be proud of this master piece by studying and understanding Thai identity, and promoting Thai art as the nation's treasure forever. This article presents a

development of Mobile device augmented reality tour application for Ramayana story on Wat Prakaew wall, which promotes the Ramayana stories on Wat Phra Kaew's wall. The user takes a photo of the wall with the application on a smartphone or a tablet connecting to the application's server via the internet. When the photo is matched with one in the database, the server sends the corresponding video clip back to be played on the mobile device. This application applies an image recognition technique and augmented reality technique to present the multimedia with more attraction. The experiment results show that the matching algorithm with extracting the painting's descriptors in the database in advance improves the processing latency by a factor of 6X. Additionally, the matching accuracy and the system response time depend on the mobile device's camera specification.

Keywords: Mobile Application, Wat Phra Kaew, Tour, Image recognition, Augmented Reality.

1. บทนำ

วัดพระศรีรัตนศาสดาราม หรือ วัดพระแก้ว ถูกสร้างขึ้นในรัชสมัยของพระบาทสมเด็จพระพุทธยอดฟ้าจุฬาโลก พร้อมกับ การสถาปนากรุงรัตนโกสินทร์ ซึ่งหลังจากสถาปนากรุงรัตนโกสินทร์ได้ 3 ปี รัชกาลที่ 1 โปรดให้เขียนภาพจิตรกรรมฝาผนังเรื่องรามเกียรติ์ขึ้น รอบพระระเบียงวัดพระแก้ว เพื่อถวายพระพุทธรูปหามณีรัตนปฎิมากร (พระแก้วมรกต) เป็นพุทธ

บุชา การดำเนินเรื่องสอดคล้องตามพระราชนิพนธ์เรื่อง รามเกียรติ์ ที่โปรดให้นักปราชญ์ราชกวี่ร่วมกันแต่งขึ้นไว้ สำหรับเป็นวรรณคดีของแผ่นดิน ชาวไทยโดยเฉพาะเยาวชนทั้งหลาย ควรภาคภูมิใจในจิตรกรรมฝาผนังนี้ ให้ความสนใจ รักและหวงแหน โดย ช่วยกันศึกษาเรียนรู้ เพื่อเข้าใจความเป็นไทย ร่วมกันสร้างสรรค์ทำนุบำรุงศิลปกรรมไทยให้เป็นศิลปะเอกลักษณ์ของชาติสืบไป

ปัจจุบันมีการพัฒนาแอปพลิเคชันบนโทรศัพท์เคลื่อนที่มากมาย เพื่อรองรับความต้องการอันหลากหลายในด้านต่างๆ ที่ให้ทั้งประโยชน์ในการทำงานและช่วยอำนวยความสะดวกต่างๆ ผู้วิจัยจึงเกิดเป็นแรงบันดาลใจในการพัฒนาโปรแกรมประยุกต์บน โทรศัพท์เคลื่อนที่เพื่อนำชมเรื่องรามเกียรติ์บนกำแพงวัดพระแก้ว ทั้งบนระบบปฏิบัติการ Android และ iOS ขึ้น

โปรแกรมประยุกต์บน โทรศัพท์เคลื่อนที่เพื่อนำชมเรื่อง รามเกียรติ์บนกำแพงวัดพระแก้วจะนำสื่อผสมมาแสดงบน โทรศัพท์ เพื่อแสดงเรื่องราวเกี่ยวกับภาพต่างๆบนฝาผนังวัดพระแก้วซึ่งเป็นรูปเกี่ยวกับรามเกียรติ์ โดยระบบจะเริ่มค้นจากการที่ผู้ใช้นำโทรศัพท์เคลื่อนที่หรือแท็บเล็ตมาถ่ายภาพบนฝาผนังวัด ซึ่งอุปกรณ์ติดตั้งระบบปฏิบัติการ Android หรือ iOS ก็ได้ และได้ติดตั้งโปรแกรมนี้ไว้แล้ว พร้อมทั้งเชื่อมต่ออินเทอร์เน็ต หากภาพถ่ายตรงกับภาพในฐานข้อมูล เครื่องแม่ข่ายก็จะส่งวีดิโอมาแสดงบนอุปกรณ์เคลื่อนที่ โดยแสดงผลทับซ้อนกับภาพบนผนังวัดที่ถ่ายจากกล้องของอุปกรณ์เคลื่อนที่แบบเวลาจริง ซึ่งวีดิโอที่แสดงนั้นจะมีความสัมพันธ์กับภาพบนฝาผนัง อีกส่วนหนึ่ง ผู้ดูแลระบบสามารถปรับเปลี่ยนภาพและวีดิโอ ในฐานข้อมูลได้ผ่านทางเว็บ

โปรแกรมประยุกต์บน โทรศัพท์เคลื่อนที่เพื่อนำชมเรื่อง รามเกียรติ์บนกำแพงวัดพระแก้วนี้ เป็นการประยุกต์ใช้เทคนิคการประมวลผลภาพ (image processing)[1] และ การเสริมความจริง (Augmented Reality) ในการนำเสนอผลงานสื่อผสม(multimedia) บนระบบปฏิบัติการ Android และ iOS โดยใช้ SURF (Speeded Up Robust Features) [2] เป็น feature detector เพื่อหา descriptor และใช้ bfmatcher เป็น ฟังก์ชันในการจับคู่ keypoints เพื่อบ่งบอกว่าภาพ 2 ภาพ เป็นภาพเดียวกัน ซึ่งเป็นไลบรารีของ OpenCV[3]

ถึงแม้ว่าจะมีโปรแกรมประยุกต์บนอุปกรณ์เคลื่อนที่สำหรับการเดินทางท่องเที่ยวที่พัฒนาขึ้นโดยใช้เทคโนโลยีเสริมความจริงในท้องตลาด แต่ยังไม่มียุคใดที่ใช้การแสดงผลแบบเสริมความจริงด้วยการนำวีดิโอคลิปมาแสดงทับซ้อนกับภาพที่ได้จากกล้องโทรศัพท์แบบเวลาจริง [4-5] ดังนั้น โปรแกรมที่ผู้วิจัยได้พัฒนาขึ้นนี้เป็น โปรแกรมแรกที่ได้ใช้เทคโนโลยีการแสดงผลแบบเสริมความจริงด้วยวีดิโอคลิปมาใช้ในการให้ข้อมูลแก่สถานที่ท่องเที่ยวในประเทศไทย

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การประมวลผลภาพ(Image Processing)

การประมวลผลข้อมูลภาพ [1] สามารถจำแนกตามวัตถุประสงค์ในการใช้งานดังนี้

- การได้มาของภาพ (Image Acquisition) โดยนำเข้าจากเซนเซอร์ และแปลงสัญญาณภาพให้เป็นภาพถ่ายดิจิทัล โดยจะมีการสุ่มตัวอย่าง (Sampling) สัญญาณภาพที่ได้รับทั้งทางพื้นที่และทางเวลา จากนั้นนำค่าที่ได้ในแต่ละพิกเซล (Pixel) มาจัดระดับข้อมูล (Quantize) เพื่อแปลงให้เป็นค่าจำนวนเต็มไม่ต่อเนื่อง (Discrete) ในรูปเลขฐานสอง เพื่อรักษาคุณภาพของภาพ จะต้องมีการดำเนินการในลักษณะที่ลดความผิดเพี้ยนที่เกิดจากเซนเซอร์ในขั้นตอนการสุ่มตัวอย่างและจัดระดับข้อมูลนี้ด้วย

- การปรับปรุงคุณภาพของภาพ (Image Enhancement and Restoration) เช่น ปรับภาพให้คมชัดมากขึ้นการปรับ contrast หรือการปรับเน้นเส้นขอบภาพ (Edge Enhancement) หรือการกรองสัญญาณรบกวน (Image Filtering) เพื่อกำจัดสัญญาณรบกวนได้

- การบีบอัดภาพ (Image Compression) เนื่องจากข้อมูลภาพนั้นมีขนาดใหญ่มาก โดยเฉพาะภาพสี ซึ่งจะทำการจัดเก็บหรือรับส่งไฟล์ข้อมูลขนาดใหญ่เสียเวลามาก เพื่อให้ขนาดภาพเล็กลงจึงทำการบีบอัดข้อมูลภาพ เช่น JPEG หรือ GIF และตัวอย่างการบีบอัด ข้อมูลวีดิโอ เช่น MPEG หรือ AVI เป็นต้น

- การวิเคราะห์ข้อมูลภาพ (Image Analysis) เป็นการสังเคราะห์ข้อมูลที่มีความหมายจากภาพแล้วนำไปใช้งาน ซึ่งอาจต้องมีรู้จำวัตถุ (object recognition) การแยกส่วนวัตถุ

(segmentation) ตรวจจับการเคลื่อนที่ของวัตถุ (motion detection)

2.2 Speeded-up robust features (SURF)

Speeded-Up Robust Features (SURF)[2] เป็นขั้นตอนวิธีที่การประมวลผลภาพ ที่ใช้ในการค้นหาจุดที่ตรงกันระหว่างสองภาพที่ เพื่อให้คอมพิวเตอร์สามารถเข้าใจหรือแยกแยะวัตถุในภาพได้ สามารถแบ่งออกเป็น 3 ขั้นตอนหลัก คือ (1) การหาจุดสนใจ โดยจะเลือกจุดที่โดดเด่นในภาพ เช่น มุม, blobs(หยดสีหรือรอยเปื้อน) และ T-junctions(จุดเชื่อมต่อรูปตัวที) โดยใช้ Hessian matrix ในการประมาณค่า และใช้ integral image ในการหาความสัมพันธ์ของแต่ละค่าข้อมูลในภาพ ทำให้การตรวจจับภาพไม่ขึ้นกับขนาด ตำแหน่ง หรือองศาการวางตัวของวัตถุในภาพ (2) หลังจากนั้น SURF จะสร้างเวกเตอร์พื้นที่โดยรอบจุดสนใจ แล้วสกัด SURF descriptor (ตัวบอกคุณลักษณะภาพ)ออกมา และ (3) จับคู่เวกเตอร์ descriptor เพื่อเปรียบเทียบภาพ

2.3 OpenCV

OpenCV(Open Source Computer Vision Library) [3] เป็นไลบรารีที่ช่วยในการเขียน โปรแกรมเกี่ยวกับ computer vision แบบเวลาจริง โดยมีการทำงานเกี่ยวกับการประมวลผลภาพแบบต่าง ๆ ที่ช่วยในการจับภาพในหลายลักษณะ เช่น การตรวจจับใบหน้า หรือสิ่งของที่ต้องการ ผ่านทางกล้อง มือถือเว็บแคม เป็นต้น OpenCV สามารถใช้ได้กับระบบปฏิบัติการ Windows , Linux , Mac , iOS และ Android โดยปัจจุบันมีขั้นตอนวิธีที่ผ่านการปรับปรุงประสิทธิภาพแล้วกว่า 2,500 ขั้นตอนวิธีให้เลือกใช้งาน โดย OpenCV รองรับภาษา C , C++ Java และ Python

2.4 ผลงานที่เกี่ยวข้อง

โปรแกรมประยุกต์บน โทรศัพท์เคลื่อนที่สำหรับการเดินทางท่องเที่ยวที่พัฒนาขึ้นโดยใช้เทคโนโลยีการเสริมความจริง (Augmented Reality) ในท้องตลาด [4-5] ได้แก่ Wikitude, junaio Augmented Reality และ Street Lens โปรแกรมทั้งสามเป็น โปรแกรมแนะนำสถานที่ โดยนำภาพสถานที่จากกล้องบน โทรศัพท์ และพิกัดผู้ใช้ มาประมวลผลร่วมกับข้อมูลแผนที่ แล้วแสดงผลโดยการนำภาพจริงจากกล้องบน โทรศัพท์เคลื่อนที่ มาซ้อนด้วยป้ายแนะนำสถานที่ (label) พร้อมระยะทางจากพิกัดปัจจุบันของผู้ใช้ โดยผู้ใช้สามารถเลือก

ประเภทของสถานที่ที่ต้องการแสดงผล และเลือกดูรายละเอียดของสถานที่ในรูปแบบของภาพและตัวอักษร

อย่างไรก็ตาม ยังไม่มีโปรแกรมนำเที่ยวใดที่ใช้การแสดงผลแบบเสริมความจริงด้วยการนำวิดีโอคลิปมาแสดงทับซ้อนกับภาพที่ได้จากกล้อง โทรศัพท์แบบเวลาจริง โดยเฉพาะอย่างยิ่ง โปรแกรมเพื่อการท่องเที่ยวในประเทศไทย ดังนั้น โปรแกรมที่ผู้วิจัย ได้พัฒนาขึ้นนี้เป็น โปรแกรมแรกที่ได้ใช้เทคโนโลยีการแสดงผลแบบเสริมความจริงด้วยวิดีโอคลิปมาใช้ในการให้ข้อมูลแก่สถานที่ท่องเที่ยวในประเทศไทย

3. การดำเนินการวิจัย

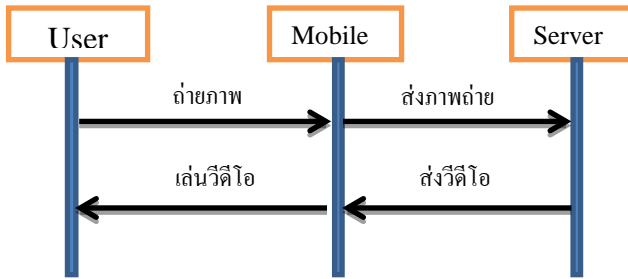
โปรแกรมประยุกต์บน โทรศัพท์เคลื่อนที่เพื่อนำชมเรื่องราวเกียรติบนกำแพงวัดพระแก้วมีการออกแบบและดำเนินการวิจัยดังต่อไปนี้

3.1 ภาพรวมของระบบ



ภาพที่ 1: สถาปัตยกรรมฮาร์ดแวร์ของระบบ

ภาพที่ 1 แสดงสถาปัตยกรรมฮาร์ดแวร์ของระบบที่พัฒนาขึ้น โดยระบบจะเริ่มต้นจากการที่ผู้ใช้นำมือถือหรือแท็บเล็ตที่เป็นระบบปฏิบัติการ Android หรือ iPhone ที่ทำงานด้วย iOS และติดตั้งแอปพลิเคชันพร้อมทั้งเชื่อมต่ออินเทอร์เน็ต มาถ่ายภาพบนฝาผนังกำแพงวัดพระแก้ว แล้วส่งไปยังเครื่องแม่ข่ายผ่านอินเทอร์เน็ต หากภาพที่ถ่ายตรงกับภาพในฐานข้อมูลเครื่องแม่ข่ายก็จะส่งวิดีโอรวมเกียรติคือนั้น ๆ มาแสดงบนมือถือ ซึ่งวิดีโอที่แสดงนั้นจะมีความสัมพันธ์กับภาพบนฝาผนัง และทับซ้อนกับภาพจริงที่ถ่ายจากกล้อง โทรศัพท์ ณ ขณะนั้น ซึ่งสามารถสรุปปฏิสัมพันธ์ระหว่าง โทรศัพท์เคลื่อนที่และเครื่องแม่ข่ายได้ดังภาพที่ 2



ภาพที่ 2: แผนภาพลำดับขั้นตอนการทำงานของระบบ

ทั้งนี้ โปรแกรมยังให้ผู้ใช้เข้ามาดูรายละเอียดหรือดาวน์โหลดวิดีโอและข้อมูลเกี่ยวกับวัดพระแก้ว แผนที่ เนื้อเรื่องรามเกียรติ์ฉบับย่อได้อีกด้วย

ในฝั่งของเครื่องแม่ข่าย ผู้ดูแลระบบสามารถปรับเปลี่ยนภาพ และวิดีโอในฐานข้อมูลได้ ผ่านทางแอปพลิเคชันบนเว็บที่ได้พัฒนาขึ้นด้วยภาษา PHP และระบบจัดการฐานข้อมูล MySQL

3.2 การเปรียบเทียบภาพถ่ายกับภาพในฐานข้อมูล

ผู้วิจัยได้ใช้ไลบรารีจาก OpenCV มาช่วยในการพัฒนา โดยใช้ภาษา C++ โดยเริ่มการทำงานด้วยการ detect feature ด้วยการหาขั้นตอนวิธีที่เหมาะสมกับการทำงาน ภาพที่ส่งจากโทรศัพท์เคลื่อนที่ที่จะมีทั้งความละเอียดสูงและความละเอียดต่ำขึ้นอยู่กับความคมชัดของกล้องหรือการตั้งค่าการถ่ายภาพของอุปกรณ์ถ่ายภาพ เมื่อมีภาพที่ส่งจากโทรศัพท์เคลื่อนที่มายังเครื่องแม่ข่าย โปรแกรมที่เครื่องแม่ข่ายจะจับคู่ภาพถ่ายกับภาพในฐานข้อมูล ซึ่งขั้นตอนวิธีที่แม่ข่ายใช้ในการสกัด feature descriptor คือ SURF [2] เนื่องจากมีความแม่นยำในการหา descriptor ที่สูงและทำงานได้รวดเร็ว ดังแสดงในหัวข้อ 4.1

หลังจาก SURF สกัด descriptors ของภาพมาเปรียบเทียบกันเพื่อตรวจสอบแล้ว จะส่งต่อไปยัง bfmatcher(BruteForce) เพื่อทำการจับคู่กับภาพในฐานข้อมูล

หลังจากนั้น ผู้วิจัยได้ปรับปรุงขั้นตอนวิธีของโปรแกรมในเครื่องแม่ข่ายที่เปรียบเทียบภาพที่ถ่ายจากโทรศัพท์เคลื่อนที่กับภาพในฐานข้อมูล จากการทดสอบรัน โปรแกรมบนโทรศัพท์เคลื่อนที่และ การรันโปรแกรมบนเครื่องแม่ข่าย พบว่าโทรศัพท์เคลื่อนที่ซึ่งมีทรัพยากรการประมวลผลต่าง ๆ จำกัดกว่าเครื่องแม่ข่าย ดังนั้นจะทำงานได้ช้ากว่ามาก จึงได้ย้ายการตรวจจับกับจับคู่ภาพในฐานข้อมูล ไปไว้ที่เครื่องแม่ข่าย

นอกจากนี้ยังปรับขั้นตอนวิธีเพื่อให้เวลาการประมวลผลลดลงอีกด้วย

3.3 การปรับขั้นตอนวิธีเพื่อลดเวลาประมวลผล

การจับคู่ภาพแต่ละภาพจะใช้เวลานานน้อยเพียงใดขึ้นอยู่กับขนาดและความละเอียดของภาพ ซึ่งภาพที่มีขนาดและความละเอียดสูงนั้นจะใช้เวลาในการประมวลผลภาพนานกว่าภาพขนาดเล็กที่มีความละเอียดต่ำ ผู้วิจัยจึงได้พัฒนาวิธีการเพื่อให้โปรแกรมสามารถทำงานได้อย่างรวดเร็วยิ่งขึ้น นั่นคือการสกัด descriptors ออกมาจากภาพในฐานข้อมูลเก็บไว้ล่วงหน้า เมื่อโปรแกรมได้รับภาพจากโทรศัพท์เคลื่อนที่ โปรแกรมจะจับคู่ descriptors ของภาพถ่ายที่ได้รับ กับ descriptors ของภาพในฐานข้อมูลได้ทันที

3.4 ขอบเขตของระบบ

ระบบที่พัฒนาขึ้นสามารถรองรับการทำงานของโทรศัพท์เคลื่อนที่และแท็บเล็ตที่ติดตั้งระบบปฏิบัติการแอนดรอยด์เวอร์ชัน 4.1 ขึ้นไป ที่มีความละเอียดหน้าจอ 1280x720 พิกเซลขึ้นไป และโทรศัพท์ iPhone ที่ใช้ iOS6 ที่มีความละเอียดหน้าจอสูงสุด 1136x640 พิกเซล

4. ผลการดำเนินการวิจัย

หลังจากพัฒนาโปรแกรมเรียบร้อยแล้ว ผู้วิจัยได้ทดสอบระบบที่ได้พัฒนาขึ้น โดยทำการทดสอบระบบกับโทรศัพท์ 4 รุ่น ได้แก่ iMobile Q4, Samsung Galaxy Note, Samsung Galaxy S4 และ iPhone5 กับภาพกิจกรรมฝาผนังในฐานข้อมูลจำนวน 30 ภาพ

4.1 ประสิทธิภาพของ SURF เทียบกับขั้นตอนวิธีอื่น

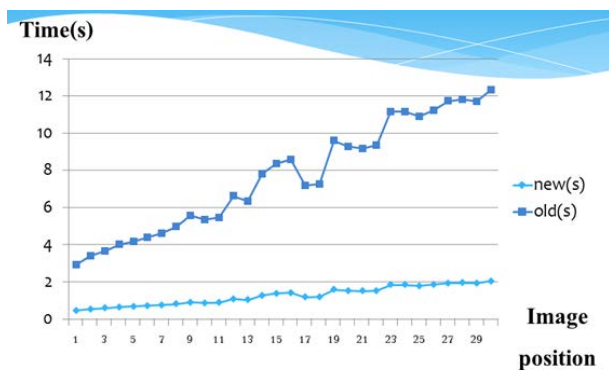
เมื่อทำการสกัด descriptors ภาพทดสอบที่มีความละเอียด 1280 x 960 พิกเซลด้วยขั้นตอนวิธี SURF, SIFT, FAST, MSER, ORB, และ DENSE ดังแสดงในตารางที่ 1 ซึ่งแสดงเวลาเฉลี่ยจากการทำงาน 5 ครั้ง พบว่า SURF สามารถสกัด descriptors ได้ปริมาณสูงกว่า FAST, MSER และ DENSE ซึ่ง descriptors ปริมาณมากเป็นผลให้ตรวจจับภาพได้แม่นยำยิ่งขึ้น แต่ SURF ยังสามารถทำงานได้เร็วกว่า ORB และ SIFT ผู้วิจัยจึงเลือกใช้ SURF ในการสกัด descriptors

ตารางที่ 1: เวลาเฉลี่ยของแต่ละขั้นตอนวิธีใช้ในการสกัด descriptors และ ปริมาณ descriptors ที่ได้

ขั้นตอนวิธี	เวลา(s)	ปริมาณ descriptor
FAST	0.1286	95bytes
MSER	0.157	95bytes
ORB	2.812	1.2MB
Dense	0.1332	96bytes
SURF	0.7846	598KB
SIFT	0.8372	1.21MB

4.2 ผลการปรับขั้นตอนวิธีเพื่อลดเวลาประมวลผล

เมื่อทำการปรับขั้นตอนวิธีของโปรแกรมบนเครื่องแม่ข่าย ดังได้อธิบายในหัวข้อ 3.3 ผลการทดสอบการจับคู่ภาพของ โปรแกรมที่ปรับเปลี่ยนแล้ว(new)และโปรแกรมที่ยังไม่ได้ปรับเปลี่ยน(old) แสดงในกราฟในภาพที่ 3 โดยแกน x แสดง หมายเลขลำดับของภาพที่จับคู่ได้ในฐานข้อมูล ส่วนแกน y คือ เวลาที่ใช้ในการประมวลผลเฉพาะ โปรแกรมส่วนการจับคู่ภาพ ที่ทำงานบนเครื่องแม่ข่าย จากกราฟพบว่าวิธีการดังกล่าว สามารถทำให้โปรแกรมประมวลผลได้เร็วขึ้นเฉลี่ย 6.00 เท่า

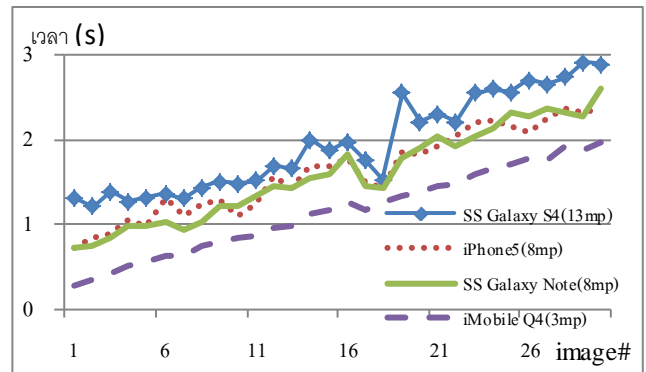


ภาพที่ 3: กราฟผลการปรับเปลี่ยนอัลกอริทึมเพื่อลดเวลาประมวลผล

4.3 ทดสอบประสิทธิภาพการทำงานโดยรวมของระบบ

ภาพที่ 4 แสดงเวลาการทำงานของระบบโดยรวม ของ โทรศัพท์ทั้ง 4 รุ่น พบว่าประสิทธิภาพขึ้นอยู่กับหมายเลขลำดับ ของภาพที่หาพบในฐานข้อมูล และความละเอียดของกล้อง โดย โทรศัพท์รุ่นที่มีกล้องที่ละเอียดกว่าจะทำงานได้ช้ากว่า เพราะ ต้องประมวลผลข้อมูลจำนวนมากกว่า ส่วนปัจจัยอื่น ไม่มี ผลกระทบมากนัก ดังแสดงในกราฟ iPhone5 และ SS Galaxy Note ที่มีความละเอียดกล้องเท่ากันและให้เวลาการทำงาน ใกล้เคียงกัน

นอกจากนี้ภาพที่ 17-18 มีการกระเพื่อมของกราฟเนื่องจาก ภาพมี descriptors จำนวนน้อย จึงใช้เวลาประมวลผลน้อยกว่า



ภาพที่ 4: กราฟเวลาการทำงานของระบบโดยรวม

4.4 ทดสอบความแม่นยำในการตรวจจับภาพ

ภาพที่ 5 แสดงผลการทดสอบความแม่นยำในการตรวจจับ ภาพของโปรแกรม บนโทรศัพท์ทั้ง 4 รุ่นที่ตั้งกล้องแบบออโต้ โฟกัส โดยความแม่นยำวัดจากจำนวนครั้งที่ถ่ายภาพแล้วเครื่อง แม่ข่ายจับคู่ภาพได้ถูกต้อง หากด้วยจำนวนครั้งที่ถ่ายภาพ ทั้งหมด(7 ครั้ง) จากผลการทดสอบพบว่าความแม่นยำขึ้นอยู่กับ

1. ความละเอียดของกล้อง
2. ความสมบูรณ์ของภาพถ่าย เช่น ภาพเต็มรูปจะให้ค่า ความแม่นยำสูงกว่าการถ่ายเพียงบางส่วนของภาพ
3. ลักษณะเฉพาะของแต่ละภาพ เช่น ลายเส้น ความแตกต่างของสี และ ตัวละครที่ปรากฏในภาพ

จากผลการทดสอบพบว่า ความแม่นยำของ โปรแกรมอยู่ที่ 100% เมื่อใช้กล้องความละเอียด 8MP ขึ้นไปและถ่ายเต็มรูป

5. สรุป

บทความนี้กล่าวถึงการพัฒนาโปรแกรมประยุกต์บน โทรศัพท์เคลื่อนที่เพื่อนำชมเรื่องราวเกียรติบนกำแพงวัดพระ แก้ว โดยโปรแกรมจะรับภาพถ่ายจากโทรศัพท์เคลื่อนที่แล้ว ส่งไปจับคู่กับภาพในฐานข้อมูลเพื่อค้นหาคลิปเรื่องราวเกียรติ ในตอนที่สัมพันธ์กับภาพมาแสดงบนหน้าจอโทรศัพท์ โดย แสดงให้ทับซ้อนกับภาพกำแพงนั้นในลักษณะเสริมความจริง โปรแกรมพัฒนาขึ้น โดยใช้ SURF ในการสกัด descriptor ของ ภาพและทำการสกัด descriptor ของภาพในฐานข้อมูลล่วงหน้า เพื่อลดเวลาการประมวลผล ซึ่งช่วยลดเวลาการประมวลผลบน เครื่องแม่ข่ายถึง 6 เท่า จากผลการทดสอบระบบพบว่า

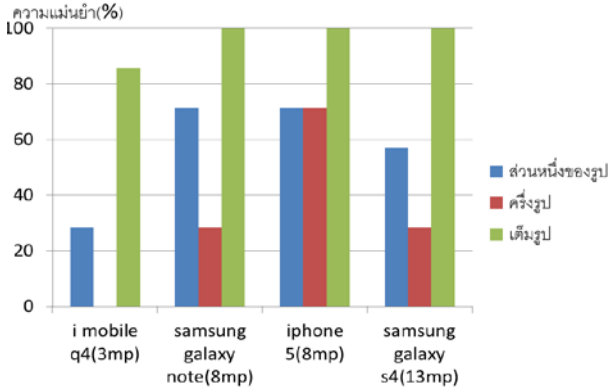
โปรแกรมสามารถทำงานได้ดีที่สุดกับโทรศัพท์ที่มีกล้องความละเอียด 8MP ขึ้นไป ที่ถ่ายภาพกำแพงเต็มรูป เนื่องจากคุณภาพของภาพมีผลต่อการสกัด descriptor ที่ใช้ในการเปรียบเทียบภาพ นอกจากนี้ เวลาที่ใช้ในการทำงานของระบบโดยรวมก็ขึ้นอยู่กับความละเอียดของกล้อง โดยภาพที่มาจากกล้องที่มีความละเอียดมากกว่าจะใช้เวลามากกว่า

ในอนาคต ต้องมีการปรับปรุงขั้นตอนวิธีในการจับคู่ภาพให้สามารถทำงานได้เร็วขึ้นอีก เนื่องจาก หากภาพที่จับคู่ถูกพบอยู่ในลำดับที่มากกว่าลำดับที่ 30 จะทำให้เวลาที่ใช้งานมากกว่า 3 วินาที ซึ่งเป็นเวลาที่นานเกินกว่าที่ผู้ใช้จะยอมรับได้ [6]

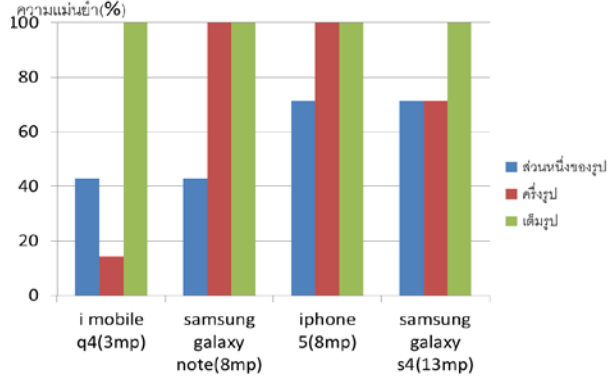
เอกสารอ้างอิง

- [1] C. Solomon and T. Breckon, *Fundamentals of Digital Image Processing: a practical approach with examples in Matlab*, Wiley-Blackwell: West Sussex, 2011
- [2] H. Bay, A. Ess, T. Tuytelaars, L. Vangool, Speeded-up robust features(surf), *Comput. Vision Image Underst.* 110 (2008) 346–359.
- [3] OpenCV Developers Team: itseez. (2013). OpenCV [Online]. Available: <http://opencv.org>.
- [4] Apple Store. <http://store.apple.com/>
- [5] Google Play. <http://play.google.com>
- [6] Kissmetrics: “How loading time affects your bottom line”. <http://goo.gl/kosva>

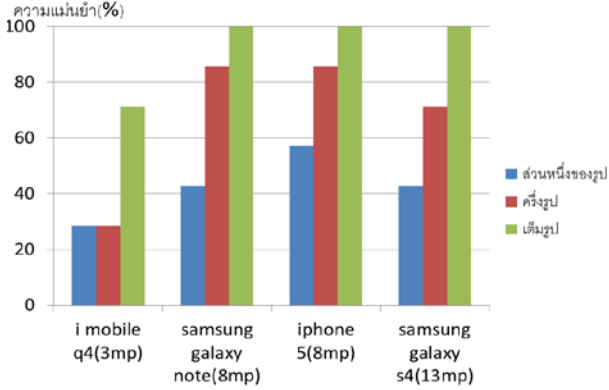
รูปห้องที่ 1 : พระชนกญาติทำพิธีบวงสรวงใต้ต้นนางสีดาแล้วลาเพศกลับเข้าเมืองมิลิดา



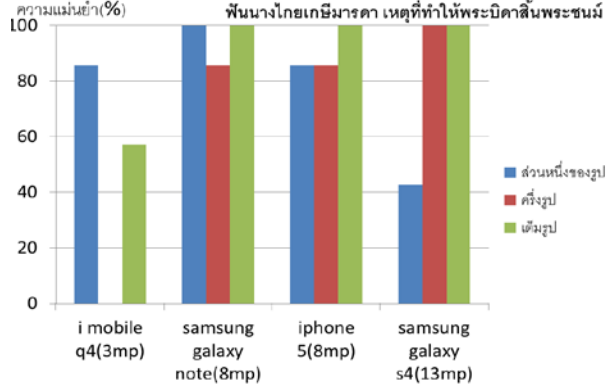
รูปห้องที่ 6 : พระพรตพระสัตรุดหลลาหัวไทยเกษเพื่อไปอยุธยา



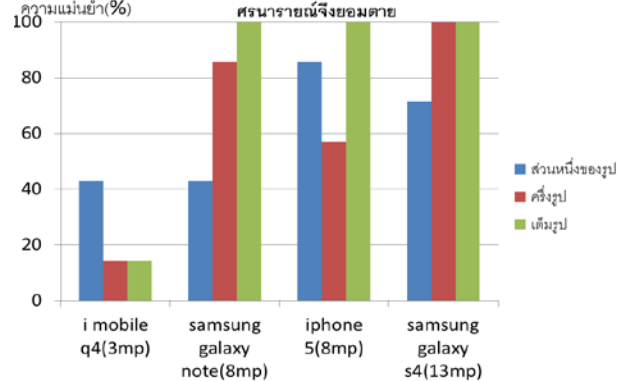
รูปห้องที่ 11 : ทรพีกำเริบทำรบพระอิศวรให้ไปทำพาลที่กรุงขีดขิน



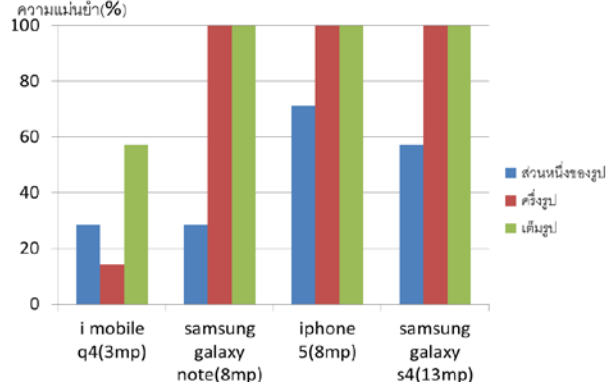
รูปห้องที่ 16 : พระรามเตรียมการบวชเป็นฤๅษี พระลักษมณ์ออกบวชตามพระพรตจะ



รูปห้องที่ 26 : หนุมานพาสุครีพมาถวายพระราม สุครีพกับพาลีรบกันจนพาลีรู้ว่าป็น



รูปห้องที่ 30 องค์กรดูหลังยักษ์ปักหลั่นให้พันสาป หนุมานส่งนางบุษมาลีขึ้นสวรรค์



**การพัฒนาโปรแกรมคอมพิวเตอร์ โดยใช้ GUI-SCILAB
ในการศึกษาการสร้างภาพเคลื่อนไหว ในการเรียนวิชาการประมวลผลภาพดิจิทัล
สำหรับนักศึกษาสาขาวิศวกรรมไฟฟ้า**

**The Development of Computer Programs using GUI-SCILAB
in a Study to Create the Image Animation in the Subject of Digital Image
Processing for Electrical Engineering Students**

กิตติ เสือแพร์¹ (Kitti Surpare)¹ และมีชัย โลหะการ² (Meechai Lohakan)²

¹ภาควิชาครุศาสตร์ไฟฟ้า คณะครุศาสตร์อุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
kittisurpare@yahoo.com, lohaganm@yahoo.com

บทคัดย่อ

การวิจัยครั้งนี้เป็นการวิจัยเชิงทดลองมีวัตถุประสงค์เพื่อพัฒนาโปรแกรมคอมพิวเตอร์โดยใช้ GUI – SCILAB ในการสร้างภาพเคลื่อนไหวสำหรับการเรียนวิชาการประมวลผลภาพดิจิทัล อีกทั้งได้ทำการหาผลสัมฤทธิ์ทางการเรียนของผู้เรียนหลังจากผ่านการใช้งานโปรแกรมดังกล่าว โดยการทดลองได้เลือกกลุ่มนักศึกษาสาขาวิศวกรรมไฟฟ้า ภาควิชาครุศาสตร์ไฟฟ้า คณะครุศาสตร์อุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ ในภาคการศึกษาที่ 2/2556 ด้วยการคัดเลือกแบบเจาะจงจำนวน 16 คน ผลการทดลองพบว่าโปรแกรมที่พัฒนาขึ้นมีประสิทธิภาพเท่ากับ 1.214 ตามสูตรของเมกุแกนส์ซึ่งสูงกว่าสมมติฐานที่ได้กำหนดไว้ นอกจากนี้ผลสัมฤทธิ์ทางการเรียนของผู้เรียนหลังจากผ่านการใช้งานโปรแกรมมีค่าสูงขึ้นอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 จึงสรุปได้ว่าโปรแกรมที่พัฒนาขึ้นสามารถนำไปใช้ประกอบการเรียนการสอนในวิชาการประมวลผลภาพได้อย่างมีประสิทธิภาพ

คำสำคัญ: การประมวลผลภาพดิจิทัล

Abstract

This research was an experimental research which purpose to develop the computer program using GUI-SCILAB. The teaching topic was the image animation in the subject of digital image processing. Additionally, the learning achievement was reported after students employ the program. The selected group was students in

Department of Electrical Engineering, Faculty of Technical Education, King Mongkut's University of Technology North Bangkok in semester 2/2556 by purposive sampling of 16 people. The results of this research found that the GUI - SCILAB was 1.214 according to Meguigan's formula, which was higher than the hypothesis. After using the program, the student's achievement was higher than before studying at the statistical significant of .01 level. In conclusion, the developed program can be used for instruction in the class efficiently.

Keyword: SCILAB, Digital Image Processing

1. บทนำ

การประมวลผลภาพได้มีการคิดค้นขึ้นในปี พ.ศ.2507 ที่ห้องทดลองของ Jet Propulsion โดยงานชิ้นแรกคือ การพิจารณาภาพถ่ายดาวเทียมของดวงจันทร์ หลังจากนั้นก็เกิดการจัดตั้งสาขาทางวิทยาศาสตร์ ชื่อว่า การประมวลผลภาพ ทำให้เกิดการสนใจในงานด้านการประมวลผลภาพเป็นอย่างมาก และใช้กันอย่างกว้างขวางสำหรับงานในหลาย ๆ ด้าน เช่น การสื่อสารทางโทรทัศน์ ทางด้านการพิมพ์ ทางด้านกราฟฟิก ทางการแพทย์ และการค้นคว้าทางวิทยาศาสตร์สำหรับในทางวิศวกรรมนั้นในปัจจุบันจะนิยมพัฒนาและทำวิจัยกันโดยใช้โปรแกรมที่มีชื่อว่า MATLAB[®]

เขาวเรศ ศิริสถิตกุล [3] ได้ศึกษาเกี่ยวกับการประยุกต์ใช้การประมวลผลภาพกับภาพจากกล้องจุลทรรศน์เพื่อวิเคราะห์คุณลักษณะส่วนประกอบโครงสร้าง และพื้นผิวภาพจุลทรรศน์เทคนิคที่ใช้สำหรับการวิเคราะห์วัตถุในภาพคือ การแบ่งส่วนภาพ (Image Segmentation) ซึ่งสามารถดำเนินการโดยใช้การกำหนดค่าเทรชโฮลด์ (Threshold Level) และการค้นหาขอบของวัตถุ (Edge Detection) ความก้าวหน้าในการวิจัยประมวลผลภาพจากกล้องจุลทรรศน์ประเภทต่าง ๆ บ่งชี้ว่า การวิเคราะห์คุณลักษณะโครงสร้างสันฐาน หรือพื้นผิวสามารถดำเนินการได้อย่างมีประสิทธิภาพ

กฤษฎา กัลยาสิริวัฒน์และ จูติ กันตถาวร [1] ได้พัฒนาโปรแกรมอ่านป้ายทะเบียนรถซึ่งเป็นเทคโนโลยีการประมวลผลภาพซึ่งใช้ในการระบุยานพาหนะโดยระบุที่ป้ายทะเบียนซึ่งพัฒนาโดยใช้ภาษาซี และ MATLAB[®] โดยโปรแกรมสามารถระบุหมายเลขทะเบียนรถได้ โดยมีค่าถูกต้องมากกว่า 50 เปอร์เซ็นต์

แม้ว่าโปรแกรม MATLAB[®] จะมีฟังก์ชันในการใช้งานสำหรับงานวิศวกรรมที่เหมาะสม แต่โปรแกรม MATLAB[®] ยังคงต้องเสียค่าลิขสิทธิ์ในการใช้งาน ซึ่งถือว่าเป็นค่าใช้จ่ายที่ค่อนข้างสูงสำหรับผู้สนใจ ไม่ว่าจะเป็น นักเรียน นักศึกษา หรือนักวิจัย

ดังนั้นในปัจจุบันจึงเริ่มมีการพัฒนางานทางวิศวกรรมขึ้นมาโดยใช้โปรแกรมที่มีความสามารถใกล้เคียงกับโปรแกรม MATLAB[®] แต่เป็นโปรแกรมฟรี ในเรื่องของค่าใช้จ่ายในการใช้งานซึ่งโปรแกรมนั้นคือ SCILAB

ในการเรียนการสอนวิชาทางด้านวิศวกรรมไฟฟ้าในเวลานี้นั้นได้มีการพยายามพัฒนาสื่อต่าง ๆ ออกมามากมาย ทั้งอยู่ในรูปของสิ่งประดิษฐ์ หรือสื่อที่ถูกพัฒนาสร้างขึ้นมาจากโปรแกรมคอมพิวเตอร์เพื่อใช้ประกอบการเรียนการสอนในวิชาต่าง ๆ

การเรียนการสอนด้านการประมวลผลภาพดิจิทัลเป็นส่วนหนึ่งของรายวิชาในหลักสูตรครุศาสตรบัณฑิต [2] สาขาวิชาวิศวกรรมไฟฟ้า ภาควิชาครุศาสตร์ไฟฟ้า คณะครุศาสตร์อุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ จากการวิเคราะห์ลักษณะเนื้อหาวิชานี้ นั้น สรุปได้ว่าลักษณะรายวิชามีเนื้อหาที่มุ่งเน้น การคำนวณทางคณิตศาสตร์

และการเขียนโปรแกรมในการจัดการเกี่ยวกับภาพซึ่งที่ผ่านมาปรากฏว่าแม้ว่าผู้สอนจะตั้งใจสอนได้ดีก็ตาม แต่ผลการเรียนของผู้เรียนยังมีความแตกต่างกันอยู่มาก ในทุก ๆ การเรียนการสอนยังคงมีผู้เรียนที่สอบตกอยู่เสมอ ซึ่งเป็นการแสดงให้เห็นว่าแม้ว่าการเรียนการสอนผู้สอนจะตั้งใจเต็มที่แล้วก็ตาม แต่ก็ยังไม่สามารถที่จะคอยตรวจปรับผู้เรียนได้ทั้งชั้นตลอดทั้งเนื้อหา เพราะถูกจำกัดด้วยปริมาณเนื้อหาและเวลาในแต่ละคาบ จึงเป็นสิ่งจำเป็นที่ผู้สอนควรมีเครื่องมือที่จะใช้ช่วยในการเรียนรู้เพื่อใช้เครื่องมือเหล่านี้เป็นตัวจัดกิจกรรมเพื่อตรวจปรับผู้เรียนในเนื้อหาที่ต้องการได้อย่างมีประสิทธิภาพ

ดังนั้นคณะผู้วิจัยจึงมีแนวคิดที่จะพัฒนาโปรแกรมคอมพิวเตอร์สำหรับประกอบการเรียนการสอนวิชาการประมวลผลภาพดิจิทัล สำหรับนักศึกษาวิศวกรรมไฟฟ้าโดยใช้ฟังก์ชัน GUI ของโปรแกรม SCILAB ซึ่งโปรแกรม SCILAB นี้มีข้อดีคือ เป็นโปรแกรมทางวิศวกรรมที่ไม่เสียค่าใช้จ่ายในการใช้งาน ทั้งยังมีศักยภาพในการทำงานทางวิศวกรรมได้ดีเทียบเท่ากับโปรแกรมอื่นๆ จึงเห็นว่าโปรแกรม SCILAB นั้นมีความเหมาะสมที่จะนำมาพัฒนาประกอบการเรียน การการคำนวณ และการเขียน โปรแกรมที่ใช้จัดการทางด้านภาพสำหรับนักศึกษาวิศวกรรมไฟฟ้าได้เป็นอย่างดี

2. วัตถุประสงค์ของการวิจัย

2.1 เพื่อพัฒนาโปรแกรมคอมพิวเตอร์ โดยใช้ GUI-SCILAB ในการศึกษาการสร้างภาพเคลื่อนไหวในการเรียนวิชาการประมวลผลภาพ

2.2 เพื่อหาประสิทธิภาพของโปรแกรมคอมพิวเตอร์โดยใช้ GUI-SCILAB ที่พัฒนาขึ้น

2.3 เพื่อหาผลสัมฤทธิ์ทางการเรียนของผู้เรียนหลังจากเรียนด้วย GUI-SCILAB ที่สร้างขึ้น

3. วรรณกรรมที่เกี่ยวข้อง

ในหัวข้อนี้ผู้วิจัยขอเสนอทฤษฎีที่นำมาใช้ในการพัฒนาโปรแกรมคอมพิวเตอร์โดยมีรายละเอียดดังนี้

3.1 โปรแกรม SCILAB

โปรแกรมภาษา SCILAB เป็นโปรแกรมที่พัฒนาโดยกลุ่มของนักวิจัยจาก INRIA และ ENPC ในประเทศฝรั่งเศสโดยมี

จุดมุ่งหมายเพื่อใช้ในการคำนวณเชิงตัวเลข ใช้ในการแสดงผล ภาพกราฟิกที่ซับซ้อน เช่นเดียวกับโปรแกรมอื่น แต่โปรแกรม SCILAB เป็นโปรแกรมที่ใช้ฟรี และมีประสิทธิภาพใกล้เคียงกับโปรแกรมอื่นเช่นโปรแกรม MATLAB ดังนั้นในปัจจุบันนี้ หลายๆ หน่วยงานทั้งภาคอุตสาหกรรม และภาคการศึกษาทั้งในและต่างประเทศได้เริ่มนำโปรแกรม SCILAB เข้ามาช่วยในการทำงานและช่วยในการเรียนการสอน ทั้งนี้เนื่องจากว่าเป็นโปรแกรมที่สามารถทำงานได้อย่างมีประสิทธิภาพ และไม่ต้องเสียเงินค่าลิขสิทธิ์ซอฟต์แวร์ และนอกจากนี้ยังมีข้อดีอื่น ๆ อีกได้แก่

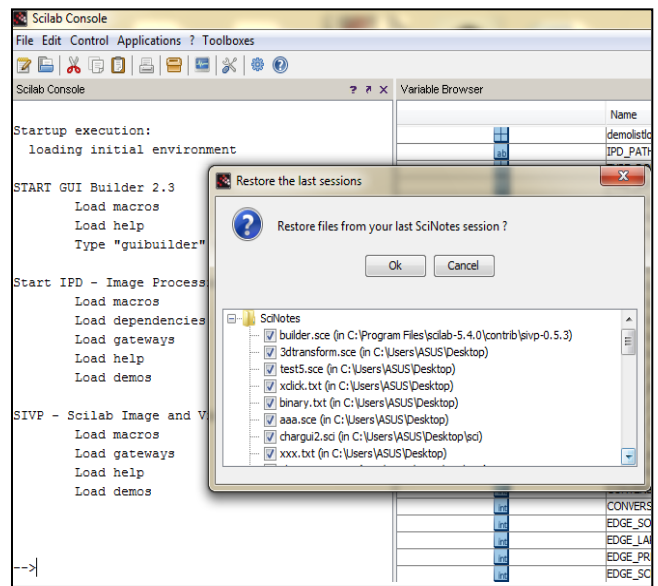
- ง่ายต่อการเรียนรู้และเข้าใจ
- ขั้นตอนการเขียนโปรแกรมไม่ยุ่งยาก
- สามารถประมวลผลข้อมูลที่อยู่ในรูปเชิงสัญลักษณ์และข้อมูลที่อยู่ในรูปเมทริกซ์ได้อย่างรวดเร็ว
- มีฟังก์ชันสำหรับการคำนวณทางคณิตศาสตร์เป็นจำนวนมาก พร้อมใช้งาน
- มีกล่องเครื่องมือจำนวนมากที่ประกอบด้วยฟังก์ชันต่าง ๆ ที่จำเป็นในการแก้ปัญหทางวิศวกรรม และวิทยาศาสตร์ หรืออื่น ๆ
- สามารถพัฒนาฟังก์ชันใหม่ ขึ้นมาใช้งานร่วมกับโปรแกรม SCILAB ได้
- สามารถใช้งานร่วมกับภาษาฟอร์แทรน ภาษาซี หรือภาษา MATLAB ได้เป็นอย่างดี
- สามารถสร้างโปรแกรมสำเร็จรูปทำซิมูเลชันระบบ
- สามารถนำไปพัฒนาต่อยอดได้เนื่องจากมี SourceCode และคู่มือการใช้งานซึ่งสามารถนำมาใช้ได้ฟรี

ซึ่งหน้าต่างของโปรแกรม SCILAB นั้นสามารถแสดงได้ดังภาพที่ 1

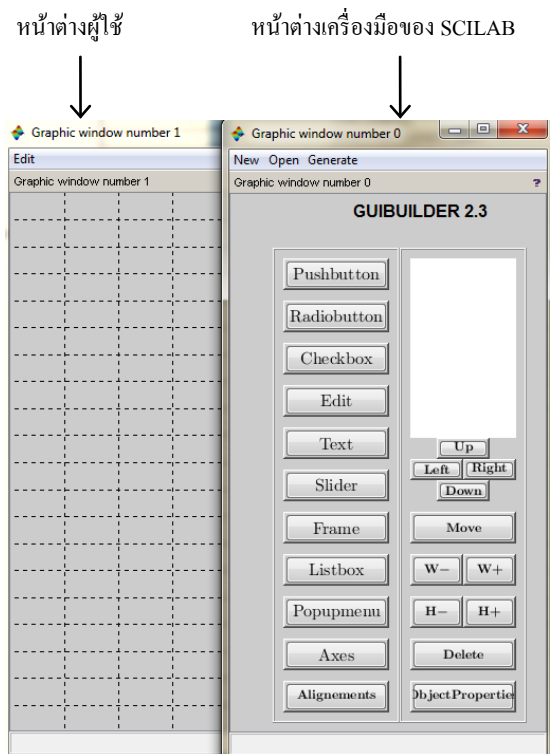
3.2 GUI

ส่วนติดต่อกับผู้ใช้ (Graphical User Interface : GUI) ได้ถูกพัฒนาขึ้นโดยนักวิจัยที่สถาบันวิจัยสแตนฟอร์ดนำโดย ดัก เอนเกลบาร์ต เป็นวิธีการใช้งานโปรแกรมคอมพิวเตอร์ผ่านทางสัญลักษณ์หรือภาพ เพื่อแทนลักษณะต่างๆ ของโปรแกรมแทนที่ผู้ใช้จะพิมพ์คำสั่งต่างๆ ในการทำงาน เพื่อช่วยให้ผู้ใช้งานสามารถทำงานได้ง่ายขึ้นและรวดเร็วขึ้น ไม่จำเป็นต้อง

จดจำคำสั่งต่างๆ ของโปรแกรมมากนัก ถือเป็นวิธีการให้ความสะดวกแก่ผู้ใช้คอมพิวเตอร์ในการติดต่อสื่อสารกับระบบคอมพิวเตอร์โดยผ่านทางภาพ



ภาพที่ 1: โปรแกรม SCILAB



ภาพที่ 2: เครื่องมือที่ใช้สร้าง GUI ของ SCILAB

ในการใช้งาน GUI ของโปรแกรม SCILAB นั้นต้องทำการติดตั้งโมดูลจากทาง ATOM ก่อนจึงจะสามารถเรียกใช้ฟังก์ชัน

สำหรับการสร้าง GUI ได้ โดยผ่านทางคำสั่ง guibuilder มีลักษณะของเครื่องมือที่ใช้สร้างดังแสดงในภาพที่ 2 ผู้ใช้สามารถลากวัตถุจากหน้าต่างเครื่องมือไปวางที่หน้าต่างผู้ใช้ โดยต้องเขียนโปรแกรมให้สอดคล้องกับวัตถุดังกล่าว

4. ขอบเขตการวิจัย

4.1 ประชากร

ประชากร คือ นักศึกษาสาขาวิชาวิศวกรรมไฟฟ้า ภาควิชา วิศวกรรมศาสตร์ไฟฟ้า มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

4.2 กลุ่มตัวอย่าง

กลุ่มตัวอย่าง คือ นักศึกษา สาขาวิศวกรรมไฟฟ้า ภาควิชา วิศวกรรมศาสตร์ไฟฟ้า มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ จำนวน 16 คน ได้มาจากการสุ่มแบบเจาะจง

4.3 ตัวแปรที่ใช้ในงานวิจัย

- ตัวแปรอิสระ คือ GUI-SCILAB เรื่องการสร้างภาพเคลื่อนไหว

- ตัวแปรตาม คือ ประสิทธิภาพของ GUI ผลสัมฤทธิ์ทางการเรียนของผู้เรียน

4.4 แบบแผนการทดลอง

การวิจัยนี้เป็นการวิจัยเชิงทดลอง ผู้วิจัยได้กำหนดแบบแผนการทดลองเป็นแบบกลุ่มเดียวสอบก่อน-สอบหลัง มีลักษณะวิธีการทดลองดังนี้

E	O ₁	X	O ₂
---	----------------	---	----------------

4.5 เครื่องมือที่ใช้ในงานวิจัย ประกอบด้วย

4.5.1 แบบทดสอบก่อน – หลังเรียน เป็นแบบทดสอบปรนัยแบบเลือกตอบสี่ตัวเลือก

4.5.2 GUI – SCILAB

4.5.3 ใบกิจกรรม

4.6 ระยะเวลาในการดำเนินงานวิจัย

วันที่ 25 มกราคม พศ. 2556

5. วิธีดำเนินการวิจัย

กระบวนการวิจัยมีวิธีดำเนินงานตามลำดับขั้นตอน ดังนี้

5.1 การวิเคราะห์เนื้อหา

ผู้วิจัยทำการวิเคราะห์เนื้อหารายวิชาการประมวลผลภาพดิจิทัล สาขาวิชาวิศวกรรมไฟฟ้า ภาควิชาวิศวกรรมศาสตร์ไฟฟ้า คณะ วิศวกรรมศาสตร์อุตสาหกรรม โดยเลือกเฉพาะเนื้อหาเรื่องการสร้างภาพเคลื่อนไหว

5.2 การออกแบบและพัฒนาโปรแกรม

ผู้วิจัยได้ออกแบบและพัฒนาโปรแกรมในส่วนการรับค่าข้อมูลอินพุต เอาท์พุตโดยใช้ GUI-SCILAB เขียนโปรแกรมให้สอดคล้องกับปุ่มต่าง ๆ ที่ได้ออกแบบไว้โดยการแสดงผลใช้กราฟในการแสดงข้อมูลเพื่อให้ง่ายในการตรวจสอบคำตอบ

5.3 การประเมินคุณภาพของโปรแกรม

นำโปรแกรมที่พัฒนาขึ้นไปให้ผู้เชี่ยวชาญจำนวน 3 ท่าน ประเมินด้านเนื้อหา และการออกแบบสื่อแล้วจึงนำมาปรับปรุงแก้ไขเพื่อไว้สำหรับการนำไปใช้ประกอบการสอน

5.4 การทดลองใช้

ทำการทดลองใช้กับนักศึกษา สาขาวิศวกรรมไฟฟ้า ภาควิชา วิศวกรรมศาสตร์ไฟฟ้า คณะ วิศวกรรมศาสตร์อุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

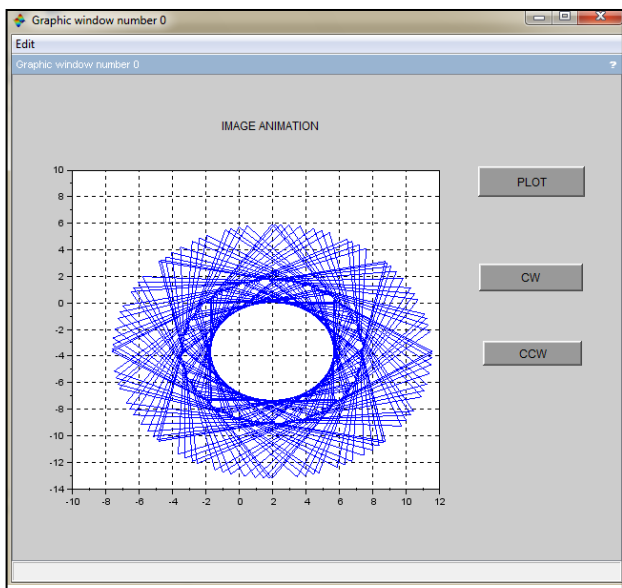
5.5 การวิเคราะห์ผล

นำ GUI-SCILAB ที่พัฒนาขึ้นและตรวจสอบโดยผู้เชี่ยวชาญเรียบร้อยแล้ว ไปใช้กับกลุ่มตัวอย่างจริง แล้วนำผลที่ได้ไปหาประสิทธิภาพของ GUI ที่พัฒนาขึ้น รวมถึงหาผลสัมฤทธิ์ทางการเรียน

6. ผลของการวิจัย

จากการดำเนินงานวิจัยจะได้โปรแกรมคอมพิวเตอร์ที่สามารถใช้ประกอบการสอนรายวิชาประมวลผลภาพดิจิทัล สำหรับนักศึกษาสาขาวิชาวิศวกรรมไฟฟ้า ภาควิชาวิศวกรรมศาสตร์ไฟฟ้า ดังแสดงในตัวอย่างดังภาพที่ 3

ผู้สอนสามารถสร้างกิจกรรมการเรียนการสอนโดยให้ผู้เรียนทำการคำนวณเปรียบเทียบกับผลลัพธ์ที่ได้จากโปรแกรม โดย ผลลัพธ์ที่ได้นี้ ผู้เรียนสามารถตรวจปรับได้ด้วยตนเองทำให้ผู้สอนใช้เวลาในการตรวจปรับผู้เรียนน้อยลง



ภาพที่ 3: GUI-SCILAB

6.1 ผลการพัฒนาโปรแกรม GUI-SCILAB

GUI-SCILAB ที่พัฒนาขึ้นประกอบด้วย แบบทดสอบก่อนเรียน แบบทดสอบหลังเรียน โปรแกรม GUI-SCILAB และใบกิจกรรม

6.2 ผลการหาประสิทธิภาพของ GUI-SCILAB

การวิเคราะห์หาประสิทธิภาพของ GUI-SCILAB เรื่องการสร้างภาพเคลื่อนไหว โดยใช้สูตรเมทริกซ์ ปรากฏผลดังตารางที่ 1

ตารางที่ 1 : ผลการวิเคราะห์หาประสิทธิภาพของบทเรียน

คะแนน	N	คะแนนเต็ม	ค่าเฉลี่ย	S.D.	ประสิทธิภาพ
ก่อน	16	15	4.5	2.1	1.214
หลัง	16	15	12.1	2.8	

จากตารางที่ 1 เมื่อนำค่าคะแนนเฉลี่ยจากแบบทดสอบก่อนเรียน กับค่าคะแนนเฉลี่ยจากแบบทดสอบหลังเรียน มาหาประสิทธิภาพตามสูตรของเมทริกซ์จะมีค่าเท่ากับ 1.214 แสดงว่ามีประสิทธิภาพตามเกณฑ์มาตรฐานของเมทริกซ์ที่มีค่ามากกว่า 1 จึงสามารถนำไปใช้ในการเรียนการสอนต่อไปได้

6.3 ผลการเปรียบเทียบผลสัมฤทธิ์ทางการเรียน

การวิเคราะห์ผลสัมฤทธิ์ทางการเรียนของผู้เรียน โดยใช้ GUI ที่พัฒนาขึ้นนี้จะนำคะแนนจากการทดสอบหลังเรียน เปรียบเทียบกับการทดสอบก่อนเรียน โดยใช้สถิติค่าที (t-test) ได้ผลดังตารางที่ 2

ตารางที่ 2 : ผลการเปรียบเทียบผลสัมฤทธิ์ทางการเรียน

คะแนน	N	คะแนนเต็ม	ค่าเฉลี่ย	S.D.	t
ก่อน	16	15	4.5	2.1	9.17*
หลัง	16	15	12.1	2.8	

จากตารางที่ 2 พบว่า ค่า t ที่คำนวณได้มีค่าเท่ากับ 9.17 เมื่อเทียบกับค่า t จากตารางที่ระดับนัยสำคัญทางสถิติ .01 และ df = 15 เปิดตารางได้ค่า t เท่ากับ 2.947 ซึ่งค่า t ที่คำนวณได้มีค่ามากกว่าค่า t ที่เปิดจากตารางจึงสรุปได้ว่าผลสัมฤทธิ์ทางการเรียนของผู้เรียนสูงขึ้นอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

7. สรุปผลการวิจัย

ผลการวิจัยพบว่า การนำ GUI-SCILAB มาใช้ในการเรียนการสอนวิชาการประมวลผลภาพ ทำให้ผู้เรียนมีการทำกิจกรรมที่ส่งเสริมให้ผู้เรียนเกิดความรู้ความเข้าใจที่มากขึ้นกว่าการเรียนทฤษฎีแต่เพียงอย่างเดียว ซึ่งการเรียนทฤษฎีแต่เพียงอย่างเดียว นั้นเป็นการที่ทำให้ผู้เรียนเกิดความเบื่อหน่ายได้ง่าย ดังนั้น GUI-SCILAB ที่พัฒนาขึ้นใช้ประกอบการเรียนการสอนนี้จะทำให้การเรียนการสอนวิชาการประมวลผลภาพมีความน่าสนใจขึ้น ดังจะเห็นได้จากค่าประสิทธิภาพตามสูตรของเมทริกซ์ซึ่งมีค่าเท่ากับ 1.214 และผลการเปรียบเทียบผลสัมฤทธิ์ทางการเรียนหลังเรียนซึ่งสูงขึ้นอย่างมีนัยสำคัญทางสถิติ ที่ระดับ .01 จึงสรุปได้ว่า GUI-SCILAB ที่พัฒนาขึ้นสามารถนำไปใช้ประกอบการเรียนการสอนวิชาการประมวลผลภาพได้อย่างเหมาะสม โดยโปรแกรมที่พัฒนาขึ้นนี้จะถูกนำไปประยุกต์ใช้ประกอบการเรียนการสอน ในวิชา การประมวลผลภาพของนักศึกษาวิศวกรรมไฟฟ้า เพื่อใช้สร้างกิจกรรมการเรียนการสอน เป็นสื่อประกอบการสอน เป็นแนวทางในการเขียนโปรแกรม และตรวจปรับความเข้าใจให้แก่นักศึกษา เพื่อประโยชน์ในการเสริมสร้างความเข้าใจให้เกิดขึ้นอย่างทั่วถึง โดยมีจุดมุ่งหมายและคาดหวังไว้ว่าจะเกิดประโยชน์สูงสุดต่อกระบวนการเรียนการสอน เพื่อให้กระบวนการเรียนการสอนวิชาการประมวลผลภาพดิจิทัลนี้สามารถจัดการเรียนการสอนได้อย่างมีประสิทธิภาพ รวมไปถึงเป็นสามารถนำไปพัฒนาต่อยอดไปสู่การสร้าง GUI ที่สามารถนำมาประยุกต์ใช้ทางด้านการประมวลผลภาพขั้นสูง ที่สามารถนำมาใช้ประโยชน์ได้ทั้งวง

การศึกษา วงการแพทย์ และงานอุตสาหกรรมต่อไปในอนาคต
อีกด้วย

เอกสารอ้างอิง

- [1] กฤษณา กัลยาสิริวัฒน์ และ จูติ กันตถาวร “โครงการระบบรู้จำป้ายทะเบียนรถ” สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหาร ลาดกระบัง
- [2] คณะครุศาสตร์อุตสาหกรรม, “หลักสูตรครุศาสตร์อุตสาหกรรมบัณฑิต (หลักสูตรปรับปรุง 2550) สาขาวิศวกรรมไฟฟ้า”, มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ กรุงเทพฯ
- [3] เยาวเรศ ศิริสถิตกุล “การประยุกต์ ใช้การประมวลผลภาพกับภาพจากกล้องจุลทรรศน์” *KKU Science Journal*. Volume 40

การพัฒนาระบบการประชุมอิเล็กทรอนิกส์ เพื่อใช้ในการประชุมกรมอนามัย กระทรวงสาธารณสุข Development of Electronic Meeting System for the Department of Health, Ministry of Public Health

พงษ์กรณ์ อักษรนิติตระกูล (Phongkorn Aksonnittrakul)¹ และศักดิ์ชัย ตั้งวรรณวิทย์ (Sakchai Tangwannawit)²
ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
¹fon589@hotmail.com, ²sakchait@kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้นำเสนอระบบการประชุมของกรมอนามัย กระทรวงสาธารณสุข ใช้เทคโนโลยีสารสนเทศที่ทันสมัยมาพัฒนาเพื่อเพิ่มความเร็วของการสื่อสาร ลดขั้นตอนการทำงานสามารถในการเปิดอ่านเอกสารได้จากทุกที่ นอกจากนี้ยังจะลดค่าใช้จ่ายในการจัดการประชุมแต่ละครั้ง เช่น ลดการใช้กระดาษ และวัสดุสำนักงาน โดยวัตถุประสงค์ของการพัฒนาระบบนี้เพื่อให้ผู้ใช้มีความสะดวกในการใช้งานซึ่งใช้หลักการของการสื่อสารระหว่างมนุษย์กับคอมพิวเตอร์ (HCI) เข้ามาช่วยในการออกแบบหน้าจอที่ใช้งานระบบ ทั้งบนสมาร์ตโฟน แท็บเล็ต และคอมพิวเตอร์ส่วนบุคคล คำนึงถึงความปลอดภัยของข้อมูลซึ่งเป็นสิ่งที่สำคัญ โดยการพิสูจน์ตัวตนผู้ใช้งานก่อนเข้าสู่ระบบเพื่อสามารถ อัปโหลดหรือ ดาวน์โหลด เอกสารตามที่ต้องการ ผลของการทดสอบในช่วง 4 เดือนนับตั้งแต่เดือนพฤศจิกายน 2013 ถึง กุมภาพันธ์ 2014 โดยการใช้งานจริงในที่ประชุมกรมอนามัย กระทรวงสาธารณสุข ผลสรุปจากการประเมินผลจากผู้ใช้งานอยู่ในเกณฑ์ที่ดี การพัฒนาระบบการประชุมผลเป็นที่ประสบความสำเร็จตามวัตถุประสงค์

คำสำคัญ: การประชุมอิเล็กทรอนิกส์ สมาร์ตโฟน แท็บเล็ต การสื่อสารระหว่างมนุษย์กับคอมพิวเตอร์

Abstract

This research proposes a conference meeting system for public health departments. This system uses modern information technology to increase the speed of communication, reduce work steps, and provide the ability to

open and read documents from anywhere. In addition, this will also decrease expenses of each meeting as well as reduce paper usage and office work. The objectives of this developed system are ease of use which includes the principles of Human-Computer Interaction (HCI) in the design of active display pages to suite both tablet smartphones and personal computers. Data security is also very important, authenticated logins are necessary before the ability to upload or download is granted. The system was tested over a period of 4 months from November 2013 to February 2014 by the Ministry of Public Health, Department of Health, Thailand. In conclusion, from the positive results of the evaluation by the Ministry of Public Health, the developed conference system clearly achieved all objectives successfully.

Keywords: e-Meeting System, Human-Computer Interaction, Tablet, Smart Phone.

1. บทนำ

กรมอนามัย ในฐานะส่วนราชการของกระทรวงสาธารณสุข ที่มีภารกิจหลักในการส่งเสริมสุขภาพ โดยส่งเสริมให้ประชาชนมีความรู้ความสามารถ และทักษะในการดูแลสุขภาพของตนเองให้ดีขึ้น และการจัดการอนามัยสิ่งแวดล้อม จึงกำหนดวิสัยทัศน์ว่า “เป็นองค์กรหลักในการส่งเสริมสุขภาพและอนามัยสิ่งแวดล้อมของประเทศ” ดังนั้น ด้วยภารกิจหลักและบทบาทหน้าที่ กรมอนามัยจึงต้องดำเนินการด้วยกระบวนการหลัก เพื่อการจัดบริการและสนับสนุนผลผลิต องค์

ความรู้ เทคโนโลยี มาตรฐาน กฎหมาย เพื่อนำไปพัฒนาให้ประชาชนตามกลุ่มวัยมีสุขภาพดี และมีสภาพแวดล้อมและสถานบริการที่ได้มาตรฐานและปลอดภัย

จากภารกิจดังกล่าว กองแผนงาน กรมอนามัย ได้รับมอบหมายในการดำเนินงานจัดประชุมผู้บริหารกรมอนามัย เพื่อรับทราบถึงนโยบาย แผนการดำเนินงาน การติดตามความก้าวหน้าการดำเนินงาน และรับทราบปัญหาอุปสรรคเป็นประจำทุกเดือน ซึ่งการจัดการประชุมในแต่ละครั้ง จะมีการใช้เอกสารประกอบการประชุมเป็นจำนวนมาก อีกทั้งงบประมาณและค่าใช้จ่ายอื่น ๆ ที่สูงเกินความจำเป็น เช่น วัสดุสำนักงาน ในการจัดทำเพิ่มเอกสาร หมึกพิมพ์ กระดาษ เป็นต้น ดังนั้นเพื่อความสะดวก รวดเร็ว ประหยัดทรัพยากรในการประชุม และขั้นตอนในการปฏิบัติงาน ผู้บริหารกรมอนามัยจึงเห็นควรให้นำระบบเทคโนโลยีสารสนเทศเข้ามาประยุกต์ใช้เพื่อแก้ปัญหาดังกล่าว

ในฐานะผู้วิจัย จึงได้เห็นโอกาสในการวิจัยหรือพัฒนาระบบดังกล่าวขึ้นโดยศึกษา วิเคราะห์ สภาพปัจจุบัน ปัญหาและความเป็นไปได้ของการนำเทคโนโลยีสารสนเทศเข้ามาช่วยเพิ่มประสิทธิภาพในการปฏิบัติงาน ทั้งในระดับปฏิบัติการและระดับบริหาร ผู้วิจัยจึงมีแนวคิดในการริเริ่มพัฒนาระบบงานนี้ คือ การจัดทาระบบการประชุมกรมอิเล็กทรอนิกส์ (e-Meeting System) สำหรับกรมอนามัย

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ค้นคว้าและศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้อง ดังนี้

2.1 ปฏิสัมพันธ์ระหว่างมนุษย์กับคอมพิวเตอร์ (HCI)

การศึกษาและวิเคราะห์ เกี่ยวกับ ปฏิสัมพันธ์ ระหว่าง มนุษย์ และ อุปกรณ์คอมพิวเตอร์ โดยมีจุดประสงค์หลัก เพื่อสร้างความสุข ความสบายให้กับผู้ใช้ เวลาที่ใช้งาน โดยคาดหวังว่าการที่ ผู้ใช้มีความสุข น่าจะทำให้มีผลผลิตที่เพิ่มขึ้น สิ่งที่จะได้จาก HCI คือ ระบบที่ ใช้ง่าย (usable) ใช้ได้ทุกคน(accessible) และมีรรถประโยชน์ (useful) ซึ่ง HCI จะครอบคลุมในเกือบทุกๆเฟส ของการพัฒนาาระบบตาม SDLC (Software Development Life Cycle) ตั้งแต่ความต้องการจากผู้บริหาร (Requirement) รวบรวมข้อมูล (Gathering) การออกแบบ

(Design) การนำไปใช้งาน (Implement) ทั้งสิ้นล้วนมี HCI เกี่ยวข้อง [1]

2.2 งานวิจัยที่เกี่ยวข้อง

ชุติมา [2] ได้ทำการศึกษาวิจัยเกี่ยวกับการพัฒนาระบบการประชุมคณะรัฐมนตรีแบบอิเล็กทรอนิกส์เพื่อใช้ในการประชุมคณะรัฐมนตรี โดยใช้พัฒนาระบบนี้เพื่อรวบรวมเอกสาร ลดขั้นตอนการจัดทำเอกสาร และส่งวารสาร ซึ่งผลการวิจัยพบว่า ความเข้าใจและการยอมรับเกี่ยวกับระบบการประชุมคณะรัฐมนตรีแบบอิเล็กทรอนิกส์ อยู่ในระดับปานกลาง ปัจจัยหนึ่งในการพัฒนา คือจะต้องวิเคราะห์ความรู้ความสามารถในการใช้สื่อสารสนเทศ ของท่านมีความรู้ความสามารถไม่เท่ากัน เป็นต้น

อดิศา [3] ได้พัฒนาระบบการประชุมอิเล็กทรอนิกส์สำหรับกรมการจัดหางาน โดยมีวัตถุประสงค์เพื่อเพิ่มประสิทธิภาพการจัดการประชุมให้มีความสะดวก รวดเร็ว ลดทรัพยากร และติดตามข้อมูลข่าวสาร ได้ทันต่อเหตุการณ์ ผลการประเมินความพึงพอใจของผู้ทดลองใช้ ทั้งด้านความสะดวกต่อการใช้งาน หน้าตาการใช้งาน โปรแกรมเข้าใจง่าย ความถูกต้องของข้อมูล โดยรวมทั้งหมดอยู่ในระดับดี

สำนักคอมพิวเตอร์มหาวิทยาลัยศรีนครินทรวิโรฒ [4] ระบบที่พัฒนาขึ้นเพื่อให้สามารถรองรับการบริหารจัดการงานประชุมของมหาวิทยาลัย โดยใช้ในงานประชุมโดยทางงานในรูปแบบของ Web Application เพื่อให้ระบบทำงานได้สะดวกและรวดเร็วขึ้น มีการเก็บสถิติและสรุปข้อมูลเกี่ยวกับการประชุม ซึ่งการพัฒนาในครั้งนี้ใช้ภาษา PHP JavaScript และใช้ MySQL ในการจัดเก็บฐานข้อมูล

3. วิธีการดำเนินการวิจัย

3.1 ศึกษาวิเคราะห์สภาพปัญหา

โดยผู้วิจัยได้เก็บปัญหาจากผู้ใช้งาน แบ่งเป็น 3 กลุ่ม คือ

3.1.1 เลขานุการการประชุม เป็นผู้ที่คอยติดตามเอกสารการประชุมในแต่ละวาระและประสานงานกับผู้ที่เกี่ยวข้องกับการประชุม โดยปัญหาสรุปได้ดังนี้

1) การติดตามเอกสารในแต่ละวาระการประชุม ซึ่งจะได้อ่อนช้ากว่า จะต้องโทรติดตามกับผู้ที่จะนำเสนอในวาระ นั้น ๆ เพื่อขอจัดเตรียมก่อนการประชุม

2) การจัดทำหนังสือการประชุม ซึ่งใช้เวลาในการเข้ารูปเล่ม และใช้ทรัพยากรค่อนข้างมาก ทั้งกระดาษ และวัสดุสำนักงานอื่นๆ เพื่อให้พอกับจำนวนผู้เข้าร่วมประชุม

3.1.2 ผู้นำเข้าข้อมูล โดยจะเป็นเลขของผู้นำเสนอวาระในที่ประชุมเป็นกรรวบรวมข้อมูล และส่งเอกสารให้กับเลขานุการการประชุม ซึ่งสรุปปัญหาได้ดังนี้

1) ข้อจำกัดทาง Email หากเป็นไฟล์ที่มีขนาดที่ใหญ่ทำให้ไม่สามารถส่งข้อมูลทางได้ จึงต้องส่งทาง CD หรือ ยูเอสบี แฟลชไดฟ์ นำข้อมูลมาให้กับเลขานุการการประชุม

2) ปัญหาทางด้านอินเทอร์เน็ต ซึ่งในบางครั้งมีความล่าช้า หรือไม่สามารถใช้งานได้ จึงทำให้เกิดความล่าช้าในการส่งข้อมูล

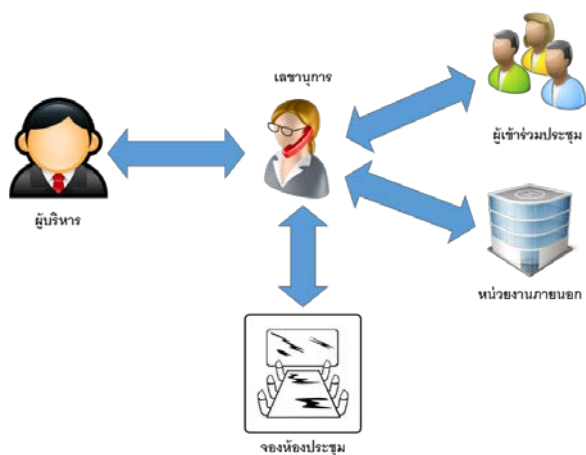
3.1.3 ผู้เข้าร่วมประชุม โดยสรุปปัญหาได้ดังนี้

1) เอกสารที่แจกเป็นรูปเล่มนั้น หากไม่ใช่ผู้บริหารระดับสูงจะเป็นรูปแบบขาว-ดำ ซึ่งในบางรายงานจะไม่สามารถเห็นความแตกต่างของข้อมูล หรือทำให้ดูค่อนข้างลำบาก

2) หลังการประชุม หากผู้บริหารต้องการเอกสารทุกวาระในที่ประชุม จะต้องนำ ยูเอสบี แฟลชไดฟ์ เพื่อมาทำการขอข้อมูลดังกล่าว หรือส่งให้ทาง Email

3) หากผู้บริหารออกตรวจราชการ จะต้องนำหนังสือการประชุมพกไปด้วยเพื่อดูข้อมูลจากที่ประชุม ซึ่งพกพาไม่สะดวก

4) การจัดเก็บเอกสาร ซึ่งปัจจุบันเอกสารมีมากขึ้นเรื่อยๆ ทำให้สิ้นเปลืองพื้นที่ในการจัดเก็บ ดังภาพที่ 1



ภาพที่ 1: ระบบการประชุมปัจจุบัน

3.2 การวิเคราะห์และแนวทางการพัฒนา

ผู้วิจัยได้ทำการศึกษาความเป็นไปได้จากการค้นคว้างานวิจัยจากแหล่งข้อมูลต่าง ๆ งานวิจัยของ ทองพล [5] ได้พัฒนาระบบการจัดการประชุมวิชาการระดับชาติให้กับทางมหาวิทยาลัยพระจอมเกล้าพระนครเหนือ จึงนำมาเป็นแนวทางในการพัฒนา และวิเคราะห์ความต้องการของผู้ใช้ด้วยวิธีการสัมภาษณ์/สอบถามความต้องการจากผู้ใช้งานทั้ง 3 กลุ่ม

ดังนั้นเพื่อต้องการแก้ไขปัญหาดังกล่าวผู้วิจัยได้ทำการพัฒนาระบบการจัดการประชุม เพื่อเป็นสื่อกลางให้ผู้นำเสนอการประชุม เลขานุการการประชุม และ ผู้ใช้งานระบบสามารถติดต่อสื่อสารการนำเข้าข้อมูลได้สะดวก และมีประสิทธิภาพมากยิ่งขึ้น การใช้งานจะคำนึงถึงหลักการใช้งานบน iPad เป็นหลัก ไม่ซับซ้อนเกินไป และง่ายต่อการใช้งาน ตามความต้องการของผู้บริหาร

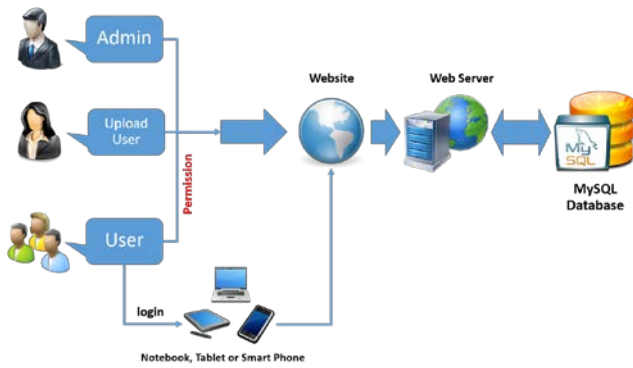
3.3 หลักการออกแบบระบบงานใหม่

การออกแบบระบบแบ่งเป็น 2 ส่วน คือ

3.3.1 ผู้ใช้งานระบบ หลังจากได้ทำการศึกษาความเป็นไปได้และการวิเคราะห์ความต้องการของระบบแล้วผู้วิจัยได้ทำการออกแบบพัฒนาระบบงานใหม่ โดยใช้งานผ่านระบบการประชุมอิเล็กทรอนิกส์กรมอนามัย และแบ่งผู้ใช้งานระบบเป็น 3 ระดับได้แก่ ผู้ดูแลฐานข้อมูล ผู้นำเข้าข้อมูล และ ผู้ใช้งานระบบ โดยระบบใหม่นี้จะช่วยลดขั้นตอนการดำเนินงานต่าง ๆ เช่น การประสานงานกับผู้นำเข้าข้อมูล โดยสามารถ นำเข้าข้อมูลเอกสารก่อนวันประชุมผ่านระบบได้โดยตรง และกำหนดเวลาในการนำข้อมูลเข้าระบบ และ ช่วยประหยัดงบประมาณ และลดการใช้อุปกรณ์สำนักงานในการจัดทำเอกสารในการประชุม ดังภาพที่ 2



ภาพที่ 2: แนวทางการพัฒนาระบบงานใหม่



ภาพที่ 3: การเชื่อมโยงการทำงานของระบบงานใหม่

3.3.2 ด้านการออกแบบต้องมีการวางแผนเพื่อตอบสนองความต้องการของผู้ใช้มากที่สุด โดยวิเคราะห์หลักการออกแบบหน้าจอการใช้งาน ดังนี้

1) ความเรียบง่าย โดยรูปแบบการใช้งานที่ง่าย คุณคุ้นเคยคล้ายกับหน้าแสดงวาระในการประชุมแต่ละครั้งที่เคยเป็นรูปแบบกระดาษ

2) ความรวดเร็วในการใช้งาน เช่น การดาวน์โหลดเอกสารจะต้องรวดเร็ว ใช้เวลาไม่นาน เพราะหากช้าเกินไป อาจทำให้ผู้ใช้งานเกิดความไม่พอใจต่อการใช้งาน ขนาดไฟล์จึงต้องบีบอัดให้มีขนาดเล็กที่สุด แต่ยังคงมีความคมชัดและมีเนื้อหาที่ครบถ้วน

3) ตัวอักษรมีขนาดใหญ่ และชัดเจน เป็นอีกปัจจัยหนึ่งที่ไม่ควรมองข้าม จากการสำรวจผู้ใช้งานระบบจะมีอายุ 40 ปีขึ้นไป หากใช้ตัวอักษรพิมพ์ที่ทั่วไป ที่ขนาด 12 point อาจทำให้ผู้ใช้งานไม่สามารถใช้งานได้สะดวกนัก ผู้วิจัยจึงเห็นว่าควรใช้ตัวอักษรที่มีขนาดใหญ่ โดยใช้ขนาด 18-20 point เพื่อให้ผู้ใช้งานสามารถอ่านได้สะดวกยิ่งขึ้น และควรเป็นตัวอักษรที่เข้ม เห็นได้ชัด ไม่ใช้สีที่สว่างจนเกินไป หรือตัวอักษรกระพริบ

4) สี เป็นสิ่งสำคัญในการทำเว็บไซต์ เพราะนอกเหนือจากข้อมูลแล้ว เฉดสีบนหน้าเว็บไซต์ ก็ถือว่าเป็นสิ่งสำคัญเป็นหน้าตาของเว็บไซต์โดยรวมเลยทีเดียว เนื่องจากสีนั้นมีอิทธิพลต่อความรู้สึกและอารมณ์ของผู้อ่าน เพราะถึงแม้ข้อมูลบนหน้าเว็บไซต์จะดีเพียงใด แต่ถ้าหากเลือกใช้โทนสีที่ไม่ดี ก็อาจส่งผลให้ดูไม่น่าสนใจหรือทำให้อ่านยาก ดังนั้น จึงควรเลือกสีให้เหมาะสมกับเว็บไซต์ว่าจะสื่อออกมาในอารมณ์และความรู้สึกแบบไหน และมี

การแสดงผลที่เอื้ออำนวยต่อการอ่านข้อมูลของผู้ชมเว็บไซต์ด้วย

ดังนั้นผู้วิจัย ได้วิเคราะห์จากกลุ่มผู้ใช้งานระบบว่ามีใครบ้าง ลักษณะงานที่ผู้ใช้ต้องการ ควรออกแบบระบบให้ตอบสนองความต้องการอย่างไร เพื่อให้ผู้ใช้ไม่ชำนาญก็สามารถใช้งานได้ไม่ยาก และการรองรับกับขนาดของหน้าจอบนอุปกรณ์ต่าง ๆ ได้ และการทดสอบก่อนการใช้งจริง

3.4 การพัฒนาระบบ

ในขั้นตอนการสร้างและพัฒนาระบบการจัดการประชุมระบบปฏิบัติการ โดยใช้การจำลองเครื่องแม่ข่าย จากเครื่องคอมพิวเตอร์ส่วนบุคคลโดยใช้โปรแกรม Apache เป็น Web Server ส่วนภาษาที่ใช้พัฒนาระบบ ได้แก่ HTML และ PHP ในการสร้างเว็บไซต์ ส่วนโปรแกรมจัดการฐานข้อมูลใช้ MySQL และในส่วนการออกแบบหน้าจอการติดต่อกับผู้ใช้งาน (Graphic User Interface) โดยใช้โปรแกรม Adobe Photoshop กับ Adobe Dreamweaver [6] ในการพัฒนา

3.5 การทดสอบประสิทธิภาพของระบบ

เมื่อทำการทดลองกับเครื่องตัวเองเป็นที่เรียบร้อยแล้ว สามารถใช้งานได้ปกติ ผู้วิจัยจึงได้ทำการนำระบบที่พัฒนาไปไว้ในระบบเครื่องแม่ข่ายจริง โดยระบบจริงที่ใช้เป็น Windows Server 2008 R2 ในส่วนของเว็บเซิร์ฟเวอร์ (Web Server) ใช้ Internet Information Server 7.5 (IIS) ที่ติดตั้งมาพร้อมกับระบบปฏิบัติการ

3.5.1 การทดสอบโดยผู้พัฒนาโปรแกรมใช้วิธีการทดสอบแบบกล่องดำ (Black Box Testing) [7] เป็นกระบวนการทดสอบการทำงานของระบบโดยรวมทั้งหมดว่ามีกระบวนการทำงานถูกต้องตามวัตถุประสงค์ที่ต้องการหรือไม่ หากพบข้อบกพร่องของโปรแกรม จะทำการปรับปรุงแก้ไขโปรแกรมให้ดีขึ้น

3.5.2 การทดสอบการแสดงผลบนอุปกรณ์ แท็บเล็ต สมาร์ทโฟน และ คอมพิวเตอร์พกพา เพื่อตรวจสอบการแสดงผลว่าเป็นไปตามที่ต้องการหรือไม่ โดยจากการทดสอบทุกอุปกรณ์สามารถแสดงผลได้อย่างที่ต้องการ ดังภาพที่ 4-6



ภาพที่ 4: การแสดงผลบน iPad และ iPhone 4/4S



ภาพที่ 5: การแสดงผลบนเครื่องคอมพิวเตอร์พกพา



ภาพที่ 6: การแสดงผลบนสมาร์ตโฟนอื่นๆ

3.5.3 การทดสอบความถูกต้องของการเชื่อมโยงข้อมูล การนำเข้าข้อมูลจากผู้นำเข้าข้อมูล ไปอยู่ในตำแหน่งจัดเก็บที่ต้องการ และป้อนเอกสาร เมื่อทำการเปิดดูจะต้องมีความถูกต้อง

ของข้อมูลในแต่ละวาระนั้น ๆ ซึ่งผลปรากฏว่าแสดงผลได้ถูกต้องตามต้องการ

3.5.4 ความรวดเร็วในการทำงานจากการทดสอบผล ปรากฏว่าระบบสามารถทำงานได้รวดเร็ว การสัมผัสหน้าจอ สามารถได้ทันที

3.5.5 การทดสอบโดยกลุ่มผู้ใช้อย่าง โดยให้ ผู้อำนวยการกองแผนงาน เลขานุการการประชุม และเจ้าหน้าที่ที่เกี่ยวข้องจำนวน 8 คน ทดสอบระบบพร้อมรับฟัง ข้อเสนอแนะเพื่อนำไปพัฒนาระบบต่อไป

4. ผลการดำเนินงาน

หลังจากเปิดใช้งานระบบจริง เป็นระยะเวลา 4 เดือน ตั้งแต่ พฤศจิกายน 2556 – กุมภาพันธ์ 2557 โดยผลการดำเนินงานและผลการวิเคราะห์ข้อมูลเพื่อหาประสิทธิภาพของระบบ ดังภาพที่ 7-9

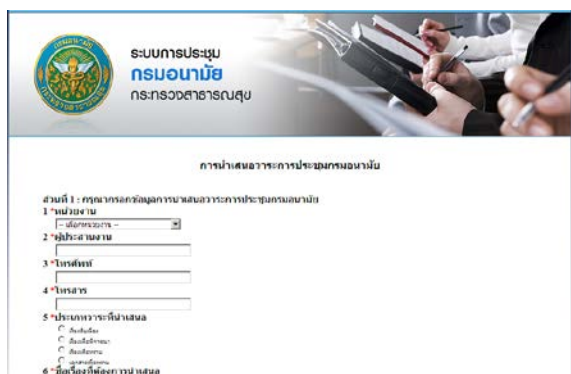
4.1 ผลการพัฒนา



ภาพที่ 7: หน้าจอพิสูจน์ตัวตนผู้ใช้งานระบบ



ภาพที่ 8: หน้าจอแสดงรายการแต่ละวาระในการประชุม



ภาพที่ 9: หน้าจอการนำเข้าข้อมูลผ่านระบบโดยผู้นำเข้าข้อมูล

4.2 ผลการวิเคราะห์และประเมินความคิดเห็นของผู้ใช้งานทั่วไป

ตารางที่ 1: ผลการวิเคราะห์แบบประเมินความคิดเห็นของผู้ใช้งานทั่วไป

รายการประเมิน	\bar{X}	S.D.	เกณฑ์
1. การประเมินด้านความสามารถทำงานตามความต้องการผู้ใช้	4.32	0.61	ระดับดี
2. การประเมินด้านหน้าที่ของโปรแกรม	4.34	0.59	ระดับดี
3. ด้านการใช้งานของโปรแกรม	4.48	0.52	ระดับดี
4. ด้านความปลอดภัยในการทำงานของระบบและข้อมูล	4.18	0.51	ระดับดี
สรุปผลการประเมินประสิทธิภาพของระบบ	4.30	0.54	ระดับดี

จากผลการประเมินความคิดเห็นของผู้ใช้งานเกี่ยวกับระบบการจัดการประชุมวิชาการผ่านเครือข่ายอินเทอร์เน็ต โดยผู้ใช้งาน จำนวน 54 คน ค่าเฉลี่ยเท่ากับ 4.30 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.54 ซึ่งสามารถแปลความได้ว่าระบบที่พัฒนาขึ้นมีประสิทธิภาพอยู่ในระดับดี

5. สรุปผลการวิจัย

5.1 สรุปผลการวิจัย

จากผลการประเมินจากการใช้ระบบการประชุม ระยะเวลา 4 เดือนจากการใช้งานจริง ตั้งแต่เดือนพฤศจิกายน 2556 ถึง กุมภาพันธ์ 2557 สามารถสรุปได้ว่าระบบที่พัฒนาขึ้นมีประสิทธิภาพการใช้งานอยู่ในระดับดี สามารถช่วยลดขั้นตอนและระยะเวลาในการดำเนินงาน อีกทั้งยังช่วยสนับสนุน

ผู้ใช้งานสามารถใช้งานได้สะดวกและรวดเร็วขึ้น ลดทรัพยากรสำนักงาน และค่าใช้จ่าย ซึ่งนำไปตามวัตถุประสงค์ที่ตั้งไว้

5.2 ข้อเสนอแนะในการพัฒนา

5.2.1 ผู้ใช้งานบางคนยังไม่คุ้นชินกับการใช้งานระบบ จึงต้องจัดทำคู่มือการใช้งาน และบริการการใช้งานกรณีติดปัญหาในที่ประชุม

5.2.2 ผู้ใช้งานมีข้อเสนอ อยากให้มีการนำระบบการประชุมแบบนี้ ไปใช้ในการประชุมอื่น ๆ ด้วย เช่น การประชุม สสอ. การประชุมวิชาการกรมอนามัย เป็นต้น

เอกสารอ้างอิง

- [1] สุริรัตน์ อินทร์หม้อ. ปฏิสัมพันธ์ระหว่างมนุษย์กับคอมพิวเตอร์. ปทุมธานี: มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี. 2552.
- [2] ชุตินา หงษ์วิเศษ. การนำระบบการประชุมคณะรัฐมนตรีแบบอิเล็กทรอนิกส์มาใช้ในการประชุมคณะรัฐมนตรี. (ออนไลน์). แหล่งที่มา: <http://goo.gl/yvDx2>. สืบค้น 24 สิงหาคม 2556
- [3] อติศา โรจนกรเกียรติ. 2556. การพัฒนาระบบการประชุมอิเล็กทรอนิกส์สำหรับกรมการจัดหางาน. (ออนไลน์). แหล่งที่มา: <http://goo.gl/ZaZUq>. สืบค้น 24 สิงหาคม 2556.
- [4] สำนักคอมพิวเตอร์ มหาวิทยาลัยศรีนครินทรวิโรฒ. ระบบการประชุมอิเล็กทรอนิกส์ มศว. กรุงเทพฯ : สำนักคอมพิวเตอร์, เมษายน 2549.
- [5] ทองพุด หีบไธสง. ระบบการจัดการประชุมวิชาการระดับชาติ. กรุงเทพมหานคร. มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ. 2551
- [6] กิตติ ภัคดีวัฒนกุล. Web Programming ด้วย Dreamweaver MX 2004 และ PHP. กรุงเทพฯ : บริษัท เลทีพี คอมพ์ แอนด์ คอนซัลท์ จำกัด. 2547
- [7] สุธิดา ชัยชมชื่น. 2556. กระบวนการ Verification & Validation เพื่อตรวจสอบและยืนยันความถูกต้องของระบบงาน. (ออนไลน์). แหล่งที่มา: ced.kmutnb.ac.th/sc/SlideASE/V_V.ppt. ตุลาคม 2556.

การพัฒนาระบบธุรกิจอัจฉริยะสำหรับสำนักอาหาร Development of Business Intelligence System for Bureau of Food

สุธีรา ปานนาค (Sutheera Phannack)¹และมาลีรัตน์ โสदानิล (Maleerat Sodanil)²

^{1,2}ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
phannack@hotmail.com¹, msn@kmutnb.ac.th²

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบธุรกิจอัจฉริยะสำหรับสำนักอาหาร สำนักงานคณะกรรมการอาหารและยา กระทรวงสาธารณสุข เป็นเครื่องมือทางด้านเทคโนโลยีสารสนเทศที่ผู้ใช้สามารถนำไปประมวลผล วิเคราะห์ ข้อมูลจำนวนมากที่มาจากแหล่งข้อมูลหลายแหล่งที่มีทั้งรูปแบบโครงสร้างข้อมูลที่มีความแตกต่างกัน เพื่อให้เป็นสารสนเทศในรูปแบบที่ผู้ใช้ต้องการได้อย่างมีประสิทธิภาพ ทำให้ผู้บริหารมีข้อมูลที่ถูกต้อง ทันสมัยและรวดเร็วผ่านเครือข่ายอินเทอร์เน็ต เพื่อใช้ประกอบการวิเคราะห์ วางแผน ติดตามผลการดำเนินงานการขออนุญาตได้อย่างมีประสิทธิภาพยิ่งขึ้น ระบบพัฒนาโดยใช้โปรแกรม Microsoft SQL Server 2008 ในรูปแบบเว็บ มีสเก็คข้อมูล แปลงรูปข้อมูลและนำไปเก็บไว้ในคลังข้อมูล ใช้เทคนิคการประมวลผลในเชิงวิเคราะห์แบบออนไลน์ (OLAP) ในการนำเสนอข้อมูลในรูปแบบของลูกบาศก์ (CUBE) ผลการดำเนินงานพบว่า จากการพัฒนาระบบธุรกิจอัจฉริยะสำหรับสำนักอาหารขึ้นมาช่วยให้ผู้บริหารได้รับข้อมูลรายงานที่รวดเร็วขึ้นและตอบสนองความต้องการของผู้ใช้งาน จากการประเมินคุณภาพระบบมีค่าเฉลี่ยอยู่ในระดับดี

คำสำคัญ: ธุรกิจอัจฉริยะ คลังข้อมูล การประมวลผลในเชิงวิเคราะห์

Abstract

The purpose of this research is to develop the business intelligence system for the Bureau of Food. It is the tool of information technology that the users can use it to process and analyze the large amount of data. These data is from the various source which have a different form and structure of data. It is a form of information which the users can efficiently use, so that the executive

have the correct and up-to-date data through the internet to support the analysis, planning, follow up the overall operation of approval for more efficient. The system used MS SQL Server 2008 as database management system and developed in web-base application style, also ETL data to data warehouse. The On-Line Analytic Processing (OLAP) technique was used to presenting data in Cube. The experiment concluded that the development of business intelligence system for the bureau of food can help the executive gets the data more quickly and timely as required. Moreover, the users are satisfied in a good level.

Keyword: BI, Data Warehouse, OLAP.

1. บทนำ

ในปัจจุบันข้อมูลของสำนักอาหารที่ถูกนำเสนอแก่ผู้บริหารองค์กรมีจำนวนมากและข้อมูลมีความหลากหลายและมีการเปลี่ยนแปลงตลอดเวลา ผู้บริหารจึงจำเป็นต้องใช้เวลาบางส่วนเพื่อใช้ในการวิเคราะห์แยกแยะหาข้อมูลที่ น่าเชื่อถือและเป็นประโยชน์ เพื่อนำไปช่วยในการตัดสินใจวางแผนกลยุทธ์ และตัดสินใจดำเนินการทางองค์กร เพื่อให้บรรลุเป้าหมายที่องค์กรวางไว้ เนื่องจากข้อมูลสารสนเทศของสำนักอาหารมีอยู่มากมายหลายแหล่ง หลายด้าน ประกอบด้วย ข้อมูลด้านคำขออนุญาตด้านอาหาร ข้อมูลด้านการตรวจสถานที่และข้อมูลอื่น ๆ ที่เกี่ยวข้อง ซึ่งข้อมูลเหล่านี้จะถูกสร้างและจัดเก็บด้วยรูปแบบและโปรแกรมที่แตกต่างกัน และจัดเก็บในแหล่งที่ต่างกัน ทำให้เกิดความยุ่งยากในการเรียกดูข้อมูลเพื่อประกอบการตัดสินใจใด ๆ ขาดประสิทธิภาพในด้านของการประมวลผลข้อมูลจากแหล่งข้อมูลที่หลากหลายและพบว่ายังมีข้อผิดพลาดอยู่มาก ไม่สามารถรองรับความ

ต้องการของผู้บริหารขององค์กรได้ และผู้ใช้เนื่องจากมีความต้องการใช้ข้อมูลนอกเหนือจากที่มีอยู่ซึ่งเจ้าหน้าที่ต้องดำเนินการจัดทำข้อมูลในรูปแบบที่ผู้บริหารต้องการเอง ทำให้ใช้ระยะเวลาในการดำเนินการมาก ไม่ทันต่อการใช้งานและอาจเกิดข้อผิดพลาดขึ้น จึงเป็นสาเหตุให้มีแนวคิดในการพัฒนาธุรกิจอัจฉริยะสำหรับสำนักอาหาร ซึ่งผู้ใช้งานแต่ละคนสามารถเลือกกำหนดเงื่อนไขต่าง ๆ ในการเข้าถึงข้อมูลได้เอง เพื่อตอบคำถาม แก้ไขปัญหาและสนองตอบความต้องการการใช้ข้อมูลสำหรับใช้ในการวิเคราะห์วางแผนการดำเนินงานหรือประกอบการตัดสินใจได้อย่างถูกต้อง สะดวก รวดเร็ว และมีประสิทธิภาพมากยิ่งขึ้น

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 Business Intelligence (BI)

Business Intelligence (BI) [1] คือ ซอฟต์แวร์ (Software) ที่นำข้อมูลที่มีอยู่เพื่อจัดทำรายงานในรูปแบบต่าง ๆ ที่เหมาะสมกับมุมมองในการวิเคราะห์ และตรงตามความต้องการของผู้ใช้งาน และใช้สำหรับวิเคราะห์ข้อมูล ของงานในมุมมองต่าง ๆ ตามแต่ละแผนก เช่นวิเคราะห์การดำเนินงานของบริษัทเพื่อการตัดสินใจด้านการลงทุนสำหรับผู้บริหาร วิเคราะห์และวางแผนการขาย การตลาด เพื่อประเมินช่องทางการจำหน่าย วิเคราะห์สินค้าที่ทำกำไร สูงสุด ขาดทุนต่ำสุด เพื่อการวางแผนงานด้านการตลาด และการผลิต วิเคราะห์ปัจจัยที่มีผลกระทบต่อยอดขายของสินค้า วิเคราะห์ข้อมูลเกี่ยวกับคู่แข่ง

2.2 คลังข้อมูล

คลังข้อมูล [2] คือ ที่เก็บข้อมูลขององค์กรที่ได้รับการออกแบบเพื่อช่วยการตัดสินใจของฝ่ายบริหารในทางปฏิบัติ นั้น สิ่งที่เกี่ยวข้องในคลังข้อมูลไม่ได้มีแต่เพียงข้อมูลเท่านั้น หากยังเก็บเครื่องมือสำหรับดำเนินการกับข้อมูล กระบวนการทำงานกับข้อมูล และทรัพยากรอื่น ๆ ระบบคลังข้อมูลเพื่อการบริหาร ได้ถูกออกแบบมาเพื่อใช้ในการเก็บข้อมูลขนาดใหญ่ โดยระบบข้อมูลเพื่อการบริหารนี้จะแยกข้อมูลออกจากฐานข้อมูลที่ใช้งานประจำวัน (Operational Database) ซึ่งข้อมูลสำหรับการบริหาร โดยมากจะเป็นข้อมูลสรุป (Summary Data) ข้อมูลสรุปนี้อาจจะเป็นข้อมูลในอดีต ข้อมูลอ้างอิง หรือข้อมูล ณ ปัจจุบัน ซึ่งอาจได้มาจากข้อมูล Operational Database หรือมีการประมวลผลข้อมูลใน Operational Database ให้

เป็นข้อมูลสรุป หรืออาจนำมาจากที่อื่นภายนอกองค์กรและทำการเพิ่มเติมลงไปก็ได้ ข้อมูลที่ถูกจัดเก็บอยู่ในคลังข้อมูลถือว่าเป็นข้อมูลในรูปแบบ Relational Database Management Systems (RDBMS) ที่มีประสิทธิภาพสูง โดยมากเราจะเลือกเก็บแต่เฉพาะข้อมูลที่จำเป็นสำหรับการตัดสินใจหรือหัวข้อของธุรกิจที่น่าสนใจ

2.3 การประมวลผลข้อมูลเชิงวิเคราะห์ (OLAP)

การประมวลผลข้อมูลเชิงวิเคราะห์ [3] คือ การใช้คำค้น (Query) เพื่อค้นหาข้อมูลในคลังข้อมูลเหมือนในฐานข้อมูล เหตุผลที่เราไม่ค้นในฐานข้อมูล แต่มาทำในคลังข้อมูลแทนมีสองสาเหตุคือ ความเร็ว ความครอบคลุมของข้อมูลทั้งบริษัทที่มีอยู่ในคลังข้อมูล ในการสำรวจข้อมูลของผู้บริหารอุตสาหกรรมส่วนใหญ่ ต้องการข้อมูล เพื่อประกอบการตัดสินใจ ในรูปแบบของสเปรดชีต (Spreadsheets) รายงาน (Report) และเครื่องมือวิเคราะห์ (Analytical tools) เครื่องมือที่สามารถนำเสนอข้อมูลและวิเคราะห์ข้อมูล ดังกล่าว คือ Online Analytical processing (OLAP) โดยโปรแกรมดังกล่าว เป็นเครื่องมือ ที่นำเสนอข้อมูลในรูปแบบต่าง ๆ จาก Data Warehouse OLAP server และ Desktop tools สนับสนุนการวิเคราะห์ความเร็วสูงของชุดข้อมูลที่มีความสัมพันธ์ที่ซับซ้อน OLAP ถูกออกแบบให้ความสามารถในการเรียกดูข้อมูล ที่รวดเร็ว เนื่องจากโครงสร้างการจัดเก็บที่เหมาะสม และการหลีกเลี่ยงการใช้ Index จำนวนมาก ทำให้ระบบแบบนี้จึงไม่เหมาะสมกับการเปลี่ยนแปลงข้อมูลจำนวนมาก ๆ ด้วยคุณสมบัติดังกล่าวนี้ ทำให้ OLAP ถูกใช้ร่วมกับ Data Warehouse เสมอ

2.4 การทบทวนวรรณกรรม

ปาณิสสา [4] นำเสนอการพัฒนาาระบบคลังข้อมูลสำหรับธุรกิจอัจฉริยะประเภทอสังหาริมทรัพย์ ผลที่ได้จากการพัฒนาจะเป็นตัวจัดเก็บข้อมูลและประมวลผลให้เป็นฐานข้อมูลเดียวกันเพื่อออกเป็นรูปแบบรายงานที่ซับซ้อนและวิเคราะห์หาคำตอบในเชิงพยากรณ์ในรูปแบบรายงานต่าง ๆ ซึ่งสามารถนำมาสนับสนุนการทำวิจัยในครั้งนี้โดยศึกษาวิธีการดำเนินการในการพัฒนาระบบธุรกิจอัจฉริยะได้

ดวงพร [5] นำเสนอระบบข้อมูลสถิติด้านการศึกษาแบบพลวัต เพื่อให้ผู้บริหารของกระทรวงศึกษาธิการมีข้อมูลที่ต้องการ ทันท่วงที สะดวก รวดเร็วผ่านเครือข่ายอินเทอร์เน็ต งานวิจัยนี้ใช้การประมวลผลข้อมูลเชิงวิเคราะห์ (OLAP) ช่วย

ในการวิเคราะห์ข้อมูลและค่าทางสถิติแบบหลายมิติ เช่นเดียวกับผู้วิจัย จึงสามารถนำมาศึกษาวิธีการดำเนินงาน ขั้นตอนการสร้างการประมวลผลข้อมูลเชิงวิเคราะห์ได้

ยวดี [6] นำเสนอการใช้เทคนิคธุรกิจอัจฉริยะในการพัฒนาระบบสารสนเทศการเข้ารับบริการงานระงับความรู้สึก อยู่ในรูปแบบเว็บแอปพลิเคชัน มีกระบวนการ ETL (Extraction, Transformation, Loading) ทำหน้าที่ดึงข้อมูลจากแหล่งต่าง ๆ ทั้งภายในและภายนอก จัดการข้อมูลให้อยู่ในรูปแบบโครงสร้างเดียวกัน เพื่อให้ข้อมูลพร้อมใช้งานได้ และนำการประมวลผลแบบออนไลน์เชิงวิเคราะห์ (OLAP) ที่ช่วยดึงข้อมูลและนำเสนอข้อมูลในหลายมิติ (Multidimensional) จากหลาย ๆ มุมมองผลการประเมินความพึงพอใจหลังจากทดลองระบบสามารถ สรุปได้ว่าระบบที่พัฒนาขึ้นสามารถนำไปใช้งานได้จริง

โชคชัย [7] ได้นำเสนอระบบคลังข้อมูลเพื่อการตัดสินใจของผู้บริหารธุรกิจการให้เข้าแบบลิสซิงสำหรับวิเคราะห์ข้อมูล และแสดงผลในรูปแบบรายงานหลายมิติ เป็นประโยชน์ในการวางแผนและกำหนดทิศทางในการพัฒนาองค์กร โดยใช้ Microsoft SQL Server ในการจัดการออกแบคลังข้อมูล การพัฒนาโปรแกรม แปลงและถ่ายโอนข้อมูล ได้การพัฒนาเครื่องมือผลิตรายงานหลายมิติที่มีความยืดหยุ่นตามความต้องการของผู้ใช้(ผู้บริหาร)

มนตรี [8] นำเสนอระบบสารสนเทศขึ้นเพื่อวิเคราะห์และประเมินสถานการณ์น้ำท่วมที่แสดงผลข้อมูลในรูปแบบธุรกิจอัจฉริยะบนเว็บแอปพลิเคชัน ใช้หลักการวิเคราะห์ในรูปแบบเจาะลึก (Drill Down) รวมทั้งการวิเคราะห์ในรูปแบบการกรองข้อมูล (Data Filtering Chart) และการวิเคราะห์แผนภูมิและแผนภูมิภาพ มีการประเมินคุณภาพโดยผู้เชี่ยวชาญใช้วิธีการแบบแบล็กบ็อกซ์ (Black Box Testing) โดยภาพรวมอยู่ในระดับที่ดี และทำการประเมินความพึงพอใจของผู้ใช้งานโดยภาพรวมอยู่ในระดับดี ดังนั้นสามารถสรุปได้ว่าระบบสารสนเทศที่พัฒนาขึ้นสามารถนำไปประยุกต์ใช้งานได้มีประสิทธิภาพ

สันติ [9] นำเสนอการวิเคราะห์ข้อมูลเพื่อการตัดสินใจในรูปแบบธุรกิจอัจฉริยะ เป็นการวิเคราะห์ถึงหลักการทำงานของธุรกิจอัจฉริยะ และความเหมาะสมของการจัดการข้อมูลเพื่อนำใช้ในการวิเคราะห์แบบออนไลน์ เพื่อประกอบการตัดสินใจที่

รวดเร็ว และมีประสิทธิภาพสามารถนำไปใช้ประกอบการดำเนินธุรกิจ โดยนำแนวคิดการใช้งานธุรกิจอัจฉริยะมาใช้ในการพัฒนาระบบแบบออนไลน์โดยสามารถแสดงผลได้อย่างรวดเร็ว

รชต [10] นำเสนอระบบสนับสนุนการตัดสินใจของผู้บริหารด้านการขาย บริษัท ไทยผลิตภัณฑ์ยิบซัม จำกัด (มหาชน) เพื่อแสดงผลการวิเคราะห์ข้อมูลใช้เป็นข้อมูลประกอบการวางแผนและและตัดสินใจดำเนินงาน โดยวิเคราะห์ข้อมูลในหลายมิติแล้วแสดงผลข้อมูลรายงานผ่านเว็บเบราว์เซอร์

บุญศรีสมิ [11] นำเสนอเทคโนโลยีคลังข้อมูลและเทคโนโลยีการประมวลผลแบบออนไลน์เชิงวิเคราะห์ (OLAP) มาพัฒนาระบบวิเคราะห์และคาดการณ์มูลค่าการค้าในลักษณะแอปพลิเคชัน โดยใช้เครื่องมือ Business Intelligence Development Studio 2008 สำหรับวิเคราะห์มูลค่าการค้าเพื่อจัดทำรายงานจากข้อมูลที่จัดเก็บอยู่ในระบบฐานข้อมูลหลายมิติ (Multidimensional Database) ทำให้เกิดความยืดหยุ่นในการออกรายงาน ระบบงานได้เพิ่มการคาดการณ์มูลค่าการค้าด้วยเทคนิคโครงข่ายประสาทเทียมแบบเพอร์เซ็ปตรอนหลายชั้นแบบการแพร่กระจายย้อนกลับ (Back Propagation) มาช่วยคาดการณ์สถานการณ์การส่งออกและนำเข้าในระยะยาวสำหรับประกอบการตัดสินใจด้านเศรษฐกิจเป็นการลดความเสี่ยงหรือขาดดุลด้านเศรษฐกิจ

สฤติชัย [12] นำเสนอระบบธุรกิจชาญฉลาดสำหรับกระบวนการด้านจดหมายติดต่อระหว่างเด็กและผู้อุปการะขององค์กรคอมแพสชั่นอินเตอร์เนชั่นแนลซึ่งใช้โปรแกรม Microsoft SQL Server 2008 ในการจัดการข้อมูล สก๊ตแปลงและบรรจุข้อมูล ส่วนของการออกรายงานใช้ Microsoft Excel 2010 Pivot Table ซึ่งสรุปผลประเมินอยู่ในระดับดี และพบว่าระบบสามารถช่วยวิเคราะห์และตอบคำถามเกี่ยวกับประสิทธิภาพการทำงานของกระบวนการเขียนจดหมาย และความสม่าเสมอของการเขียนจดหมายของเด็กได้ดี

3.วิธีการดำเนินการ

3.1 การศึกษาและรวบรวมข้อมูลระบบ

ศึกษาการดำเนินงานในปัจจุบัน พร้อมทั้งรวบรวมข้อมูลความต้องการผู้ใช้งาน ศึกษาแนวคิดการทำระบบธุรกิจ

อัจฉริยะ การทำคลังข้อมูล และขั้นตอนการพัฒนา ระบบด้วยเทคนิคการประมวลผลข้อมูลเชิงวิเคราะห์ (OLAP) ศึกษาเครื่องมือที่จะนำมาใช้ในการพัฒนาระบบ พร้อมทั้งศึกษาข้อมูลฐานข้อมูลที่มีอยู่ในปัจจุบัน

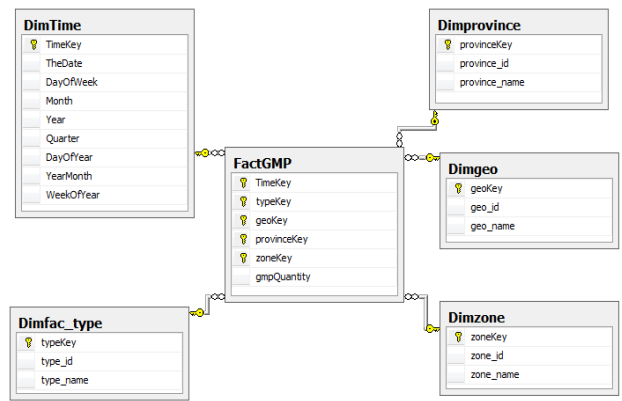
3.2 การออกแบบระบบ

หลังจากที่ได้ศึกษาการดำเนินงานปัจจุบัน วิเคราะห์ปัญหาปัจจุบัน รวบรวมความต้องการของผู้ใช้งานเรียบร้อยแล้ว ก็ได้ทำการออกแบบโครงสร้างของระบบใหม่ วิเคราะห์ข้อมูลที่จะนำมาใช้ในระบบใหม่ ซึ่งการออกแบบและพัฒนาระบบใช้ Microsoft SQL Server 2008 เป็นเครื่องมือหลัก ระบบจะนำข้อมูลเสนอผ่านเว็บไซต์ โดยใช้การประมวลผลข้อมูลเชิงวิเคราะห์ (OLAP) ช่วยในการวิเคราะห์ข้อมูลแบบหลายมิติ ผู้ใช้งานสามารถเลือกตัวแปรและค่าของข้อมูลได้ทั้งในรูปแบบตารางรายงานและแผนภูมิต่าง ๆ



ภาพที่ 1: โครงสร้างระบบ

จากภาพที่ 1 แสดงโครงสร้างของระบบ มีการนำข้อมูลที่เกี่ยวข้องในรูปแบบต่าง ๆ จากหลาย ๆ แหล่งข้อมูลมาลดความซ้ำซ้อน ทำการจัดเก็บในรูปแบบเดียวกัน โดยผ่านกระบวนการสกัด แปลง และบรรจุข้อมูล (ETL) นำมาจัดเก็บลงในคลังข้อมูล แล้ววิเคราะห์ข้อมูลในคลังข้อมูลด้วยเทคนิคการประมวลผลข้อมูลเชิงวิเคราะห์ (OLAP) เพื่อนำเสนอข้อมูลในมุมมองต่าง ๆ และแสดงรายงานสรุปแก่ผู้ใช้งาน

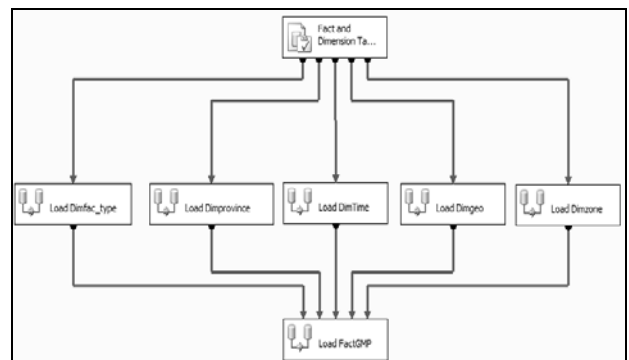


ภาพที่ 2: โครงสร้างการออกแบบตารางข้อมูลตามลักษณะของ Dimensional Database

จากภาพที่ 2 แสดงโครงสร้างการออกแบบตารางภายในคลังข้อมูลของระบบธุรกิจอัจฉริยะ โดยใช้ Microsoft SQL Server 2008 ซึ่งตารางข้อมูลต่าง ๆ มีความสัมพันธ์ระหว่างกันในลักษณะของ Dimensional Database จะประกอบด้วย Fact Table และ Dimension Table โดยใช้ Primary Key ของแต่ละ Dimension Table เชื่อมโยงไปยัง Fact Table Key

3.3 การพัฒนาระบบ

ในการพัฒนาระบบได้เลือกใช้เครื่องมือ Microsoft SQL Server 2008 ในการจัดการสร้างพัฒนาคลังข้อมูลสำหรับระบบธุรกิจอัจฉริยะตามแบบจำลองที่ได้ออกแบบไว้ รวมถึงการนำเข้าข้อมูลและการถ่ายโอนข้อมูล (ETL) จากแหล่งข้อมูลต่าง ๆ พร้อมทั้งใช้แสดงผลรายงานในรูปแบบเว็บ



ภาพที่ 3: ผังการทำงานหลักสำหรับการถ่ายโอนข้อมูล

จากภาพที่ 3 แสดงผังการทำงานหลักสำหรับการถ่ายโอนข้อมูลจากแหล่งข้อมูลที่กำหนด แปลงข้อมูลให้มีรูปแบบที่สอดคล้องกับโครงสร้างในคลังข้อมูลตาม Schema ที่ออกแบบไว้ พร้อมทั้งตรวจสอบความถูกต้องและแก้ไขข้อมูลที่จะเก็บลงคลังข้อมูลให้มีความถูกต้อง

Measure Group	Dimension	Hierarchy	Operator	Filter Expression
<All>	Dim Fac	จังหวัด	Equal	{All}
	Dim Time	ไตรมาส	Equal	{All}
		<Select dimension>		

Year	Type Name	2013	2014	Grand Total
Province Name	Fact Count	699	3333	4032
Province Name	Fact Comp Count	157	290	347
Province Name	Fact Comp Count	239	244	483
Province Name	Fact Comp Count	2	948	950
Province Name	Fact Comp Count	517	900	1417
Province Name	Fact Comp Count	19	21	40
Province Name	Fact Comp Count	10	10	20
Province Name	Fact Comp Count	18	35	53
Province Name	Fact Comp Count	11	41	52
Province Name	Fact Comp Count	17	7	24
Province Name	Fact Comp Count	8	22	30
Province Name	Fact Comp Count	22	76	98
Province Name	Fact Comp Count	105	197	302
Province Name	Fact Comp Count	2	104	106
Grand Total	Fact Count	2380	2114	4494

ภาพที่ 4: หน้าจอผลจากการออกแบบ Cube

จากภาพที่ 4 แสดงส่วนของการสร้างและแสดง Cube ตามมุมมองที่กำหนดไว้สำหรับการแสดงผลรายงาน

3.4 การทดสอบระบบ

การทดสอบโดยผู้พัฒนาระบบ ใช้วิธีการทดสอบประสิทธิภาพแบบแบล็กบ็อกซ์ (Black Box Testing) เป็นกระบวนการทดสอบการทำงานของระบบโดยรวมทั้งหมดว่ากระบวนการทำงานถูกต้องตามวัตถุประสงค์ที่ต้องการหรือไม่

การทดสอบระบบในขั้นเบต้า (Beta Stage) นำระบบไปทำการทดสอบประสิทธิภาพ โดยผู้เชี่ยวชาญจำนวน 5 คน พร้อมทำแบบประเมินคุณภาพของระบบ โดยกำหนดหัวข้อของการทดสอบคือ ประเมินด้านความสามารถของระบบตรงตามความต้องการของผู้ใช้งาน ประเมินด้านความถูกต้องของการทำงานของระบบ และประเมินด้านการใช้งานระบบ โดยกำหนดเกณฑ์ในการทดสอบประเมินไว้ 5 ระดับ ดังตารางที่ 1

ตารางที่ 1: เกณฑ์การให้คะแนนแบบประเมินความพึงพอใจ

เกณฑ์การให้คะแนน	ความหมาย
5	ระบบที่พัฒนามีความเหมาะสมมากที่สุด
4	ระบบที่พัฒนามีความเหมาะสมมาก
3	ระบบที่พัฒนามีความเหมาะสมปานกลาง
2	ระบบที่พัฒนามีความเหมาะสมน้อย
1	ระบบที่พัฒนามีความเหมาะสมน้อยที่สุด

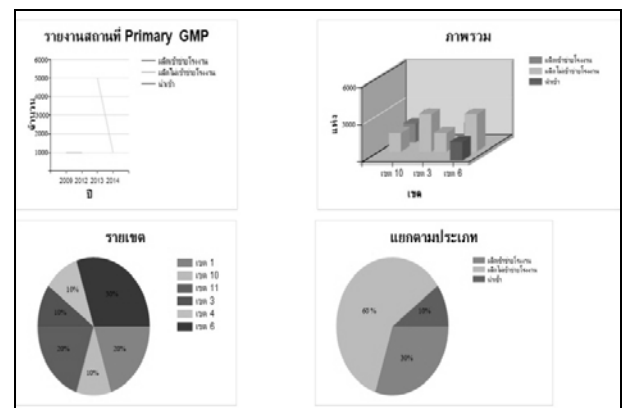
3.5 การประเมินระบบ

ขั้นตอนการสรุปผลของระบบ เมื่อระบบถูกทดสอบและประเมินคุณภาพจากผู้เชี่ยวชาญเป็นที่เรียบร้อยแล้ว จากนั้นทำการวิเคราะห์ข้อมูลทางสถิติซึ่งประกอบด้วยค่าเฉลี่ยและส่วน

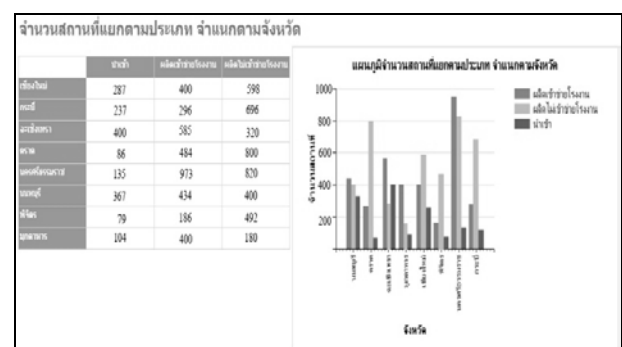
เบี่ยงเบนมาตรฐาน นำผลที่ได้มาสรุปผลและหาข้อเสนอแนะเพื่อนำไปพัฒนาปรับปรุงให้มีประสิทธิภาพมากขึ้น

4. ผลการดำเนินงาน

ผลการพัฒนาระบบ ระบบธุรกิจอัจฉริยะของสำนักอาหารสามารถแสดงผลข้อมูลจาก Dimension และ Cube ที่ออกแบบไว้ได้เหมาะสมและสอดคล้องต่อความต้องการของผู้ใช้งาน ซึ่งผู้ใช้ระบบสามารถที่จะทำการเรียกดูรายงานได้ดังนี้ รายงานสถานที่ทั้งหมด รายงานสถานที่แยกประเภท รายงานสถานที่ตามผลการตรวจ รายงานสถานที่ตามเขต รายงานสถานที่ตามภูมิภาค และรายงานแบบหลายมิติ ซึ่งถือว่าเป็นหัวใจหลักของระบบ โดยผู้ใช้สามารถที่จะทำการเรียกดูข้อมูลในแบบเจาะลึก (Drill Down) หรือแบบแสดงผล และจากผลการประเมินคุณภาพระบบจากผู้เชี่ยวชาญของสำนักอาหารสรุปว่ามีค่าเฉลี่ยอยู่ในระดับดี



ภาพที่ 5: ตัวอย่างหน้าจอ Dashboard



ภาพที่ 6: ตัวอย่างหน้าจอแสดงรายงาน 1 มิติ

		๒ 2011	๒ 2012	๒ 2013	
		ผลิตเข้าขาย โรงงาน	ผลิตเข้าขาย โรงงาน	ป่าเข้า	ผลิตไปเข้าขาย โรงงาน
๒ เขต 1	เชียงใหม่	260	140	287	598
๒ เขต 10	มุกดาหาร	124	276	104	180
๒ เขต 11		4920	7639	3389	5490
๒ เขต 3		3452	3870	2785	4432
๒ เขต 4	นนทบุรี	178	256	376	400
๒ เขต 6	ฉะเชิงเทรา	243	342	400	320
	ตราด	117	367	86	800

ภาพที่ 7: ตัวอย่างหน้าจอแสดงรายงานหลายมิติ

จากภาพที่ 5 ภาพที่ 6 และภาพที่ 7 แสดงตัวอย่างหน้า
หน้าจอ Dashboard สำหรับผู้บริหาร ตัวอย่างหน้าจอแสดง
รายงาน 1 มิติ และตัวอย่างหน้าจอแสดงรายงานหลายมิติ
ตามลำดับ โดยผู้ใช้งานระบบสามารถเรียกดูข้อมูลผ่านเว็บ
เบราว์เซอร์ในแบบเจาะลึก (Drill Down) และเรียกดูตาม
เงื่อนไขได้

5. สรุปผล

งานวิจัยครั้งนี้ได้ทำการพัฒนาระบบธุรกิจอัจฉริยะสำหรับ
สำนักอาหาร โดยงานวิจัยนี้ได้ใช้ชุดเครื่องมือ Microsoft SQL
Server 2008 ในการพัฒนา พร้อมทั้งแสดงผลในรูปแบบของ
เว็บ ผลจากการศึกษา พบว่า ระบบที่พัฒนาขึ้นสามารถสร้าง
รายงานและเรียกดูข้อมูลรายงานแบบหลายมิติที่สามารถ
เจาะลึกลงรายละเอียด (Drill Down) หรือเรียกดูข้อมูลสรุปใน
ภาพรวม (Roll Up) ได้ หลังจากได้ทำการพัฒนาระบบเสร็จสิ้น
ทำการประเมินคุณภาพระบบพบว่า ผลจากการประเมินอยู่ใน
ระดับดี

เอกสารอ้างอิง

[1] โอภาส เอี่ยมสิริวงศ์, การวิเคราะห์และออกแบบระบบ ฉบับ
ปรับปรุงเพิ่มเติม System Analysis and Design, กรุงเทพฯ,
ซีเอ็ดยูเคชั่น, 2549.

[2] กิตติพงษ์ กลมกล่อม, การออกแบบและพัฒนาคู่มือข้อมูล
Data Warehouse, กรุงเทพฯ, สำนักพิมพ์ เคทีที คอมพ์
แอนด์ คอนซัลท์, 2552.

[3] โอม ศรีนิล, การออกแบบและพัฒนาคู่มือข้อมูล, กรุงเทพฯ,
บางกอกบล็อท, 2556.

[4] ไพบุลย์ เกียรติโกมล, ระบบสารสนเทศเพื่อการจัดการ, กรุงเทพฯ,
ซีเอ็ดยูเคชั่น, 2551.

[5] ปาณิสรา วัชรสินธุ และประสงค์ ปรานิตพลกรัง, ระบบคลังข้อมูล
สำหรับธุรกิจอัจฉริยะประเภทอสังหาริมทรัพย์ของธนาคาร,
สาขาวิชาเทคโนโลยีสารสนเทศ บัณฑิตวิทยาลัย มหาวิทยาลัย
ศรีปทุม, 2552.

[6] ดวงพร เขียมอัมพร, สุพจน์ นิตย์สุวัฒน์, มนต์ชัย เทียนทอง และ
พวง มีสัง, “ระบบนำเสนอข้อมูลสถิติด้านการศึกษาแบบพลวัต
ของกระทรวงศึกษาธิการ”, *The National Conference on
Computing and Information Technology NCCIT'08*,
หน้า 633 – 638, 2551.

[7] ยูดี ฟ้าทอง, การพัฒนาระบบสารสนเทศการเข้ารับบริการ
งานระดับความรู้ลึก(วิสัยทัศน์), วิทยาศาสตร์มหาบัณฑิต
สาขาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2554.

[8] โชคชัย บุญมาลัย, ระบบคลังข้อมูลเพื่อการตัดสินใจของ
ผู้บริหารธุรกิจการให้เช่าแบบลิสซิง, สาขาวิชาวิทยาการ
คอมพิวเตอร์ บัณฑิตวิทยาลัย มหาวิทยาลัยเชียงใหม่, 2551.

[9] มนตรี มงคลทรง, ระบบวิเคราะห์ปริมาณสถานการณ์น้ำท่วม
โดยแสดงผลข้อมูลในรูปแบบธุรกิจอัจฉริยะ, วิทยาศาสตร์มหาบัณฑิต
สาขาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2555.

[10] สันติ อริยทรัพย์, เท็ดพงษ์ ผลจันทร์ และปรีชา น้อยคำภา,
“การวิเคราะห์ข้อมูลเพื่อการตัดสินใจด้วย Business Intelligence”,
*การสัมมนาวิชาการ ครั้งที่ 3 National CS & ICT Confernce
2009, มหาสารคาม, 2552.*

[11] รชต เดชาธรรพล, ระบบสนับสนุนการตัดสินใจของผู้บริหาร
บริษัท ไทยผลิตภัณฑ์อียิปต์ จำกัด (มหาชน), สาขาวิชาการ
จัดการเทคโนโลยีสารสนเทศ มหาวิทยาลัยวลัยลักษณ์, 2550.

[12] ปญุชร์สมิ์ สว่างประเสริฐ, ระบบวิเคราะห์และคาดการณ์มูลค่า
การค้าโดยเทคนิคโครงข่ายประสาทเทียม, วิทยาศาสตร์มหาบัณฑิต
สาขาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2554.

[13] สติชัยชัย ลิขิตพานิชกุล, การพัฒนาระบบธุรกิจชาญผลลาด
สำหรับกระบวนการด้านจดหมายติดต่อระหว่างเด็ก
และผู้อุปการะขององค์กรคอมพิวเตอร์เอ็นซันแนล,
สาขาวิชาเทคโนโลยีสารสนเทศและการจัดการ บัณฑิตวิทยาลัย
มหาวิทยาลัยเชียงใหม่, 2554.

การออกแบบเป้าหมายและตัวชี้วัดประสิทธิภาพการให้บริการรับชำระค่าสินค้าและบริการ กรณีศึกษาองค์กรการสื่อสารและโทรคมนาคม

The Design of Targets and Performance Indicators for Payment Service: Telecommunication Organization Case Study

วีราภรณ์ ซิดดู (Veeraporn Siddoo)¹ และ รัตนา เวทย์ประสิทธิ์ (Rattana Wetprasit)²

สาขาวิชาเทคโนโลยีสารสนเทศและการสื่อสาร คณะเทคโนโลยีและสิ่งแวดล้อม

มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตภูเก็ต

¹veeraporn.s@phuket.psu.ac.th, ²rattana.w@phuket.psu.ac.th

บทคัดย่อ

งานวิจัยนี้นำเสนอการออกแบบเป้าหมาย ตัวชี้วัด และเกณฑ์การวัดประสิทธิภาพการให้บริการรับชำระค่าสินค้าและบริการขององค์กรการสื่อสารและโทรคมนาคมแห่งหนึ่งในประเทศไทย ใช้วิธีวิจัยแบบกรณีศึกษา (Case Study Methodology) ด้วยการศึกษารายงานการทำงานของแผนกให้บริการรับชำระค่าสินค้าและบริการ รวมถึงระบบสารสนเทศที่มีอยู่ ผลการศึกษาทำให้ได้เป้าหมายประกอบตัวชี้วัดประสิทธิภาพ 5 ตัวชี้วัดพร้อมสูตรการคำนวณ ผู้วิจัยนำเสนอผลที่ได้ด้วยการเชื่อมโยงมิติ Balance Scorecard (BSC) กับเป้าหมาย และวิธีวัดความสำเร็จโดยใช้ตัวชี้วัดและเกณฑ์การให้คะแนนพร้อมอภิปรายผลการวิจัยและแนวทางการปรับปรุงประสิทธิภาพ ทำให้องค์กรที่ดำเนินธุรกิจประเภทเดียวกันสามารถนำผลงานวิจัยไปประยุกต์ใช้ได้ตามความเหมาะสม

คำสำคัญ: การวัดประสิทธิภาพ บาลานซ์สกอร์การ์ด ดัชนีตัวชี้วัด

Abstract

This paper presents the design of targets, indicators and measurement criteria to measure the payment service's performance using the Case Study Methodology. We had studied the work flow of the payment service department in a telecommunications organization in Thailand. Moreover, we investigated the existing ICT systems and verified information. The result of paper gives us five targets and indicators, including indicator calculation formulas. We demonstrated our

output by linking them to Balance Scorecard (BSC) dimensions. We showed the findings and also gave the suggestions to improve organization's performance. Organizations who do the same business can apply this case study as appropriate.

Keyword: performance measurement, balance scorecard, key performance indicator

1. บทนำ

ความต้องการสร้างความพึงพอใจให้กับลูกค้าในการใช้บริการรับชำระค่าสินค้าและบริการด้วยความรวดเร็ว ถูกต้อง และสะดวกสบาย เป็นที่มาของการจัดทำงานวิจัยขึ้นนี้ ความสำเร็จที่จะเกิดขึ้นดังกล่าวต้องมาจากองค์ประกอบหลายด้าน ทั้งจากความร่วมมือที่ดีของพนักงานองค์กรที่ให้บริการ การมีเทคโนโลยีที่ทันสมัยช่วยในการอำนวยความสะดวก หรือแม้แต่การจัดองค์ประกอบของสถานที่ให้ลูกค้าเกิดความประทับใจ นำมาสู่การมาใช้บริการในครั้งต่อไป องค์กรกรณีศึกษาให้บริการรับชำระค่าสินค้าและบริการ ครอบคลุมงานดังนี้ รับชำระค่าใช้บริการ โทรศัพท์มือถือ บัตรเครดิตต่างๆ บริการรับฝากชำระกลุ่มลิสซิ่ง กลุ่มประกันชีวิตและประกันภัย กลุ่มสินเชื่อ และกลุ่มสินค้าและบริการ โดยมีการประเมินผลการดำเนินการไม่น้อยกว่า 4 ครั้ง ต่อปี ภายในกรอบเวลาทุกๆ 3 เดือน ด้วยขั้นตอนการทำงานที่มีอยู่เดิมพบว่า ยังไม่เคยกำหนดการวัดประสิทธิภาพอย่างเป็นทางการ ทำให้ยากที่จะบอกได้ว่าบริการให้บริการรับชำระค่าสินค้าและบริการที่เป็นอยู่ประสบความสำเร็จดีแล้วหรือยังมีจุดที่ต้องปรับปรุง

กระบวนการทำงานเดิมให้ดีขึ้นอีก ระบบสารสนเทศที่มีก็เป็นการทำงานที่แยกจากกันในแต่ละระบบ ยังไม่เคยมีการนำข้อมูลมาใช้ร่วมกันเพื่อประเมินประสิทธิภาพการให้บริการ จากประสบการณ์ทำงานในองค์กรนี้ ผู้วิจัยจึงมองเห็นแนวทางในการนำข้อมูลดังกล่าวมาใช้ให้เกิดประโยชน์

งานวิจัยนี้เป็นการศึกษาขั้นตอนการทำงานรวมถึงปัญหาที่มักเกิดขึ้น มาประกอบการทำวิจัย เพื่อกำหนดเป้าหมายและสร้างตัวชี้วัดประสิทธิภาพการดำเนินงานว่าบรรลุตามเป้าหมายที่ตั้งไว้หรือไม่ การวัดประสิทธิภาพจะวัดโดยใช้ดัชนีตัวชี้วัดประสิทธิภาพ (Key Performance Indicator: KPI) ซึ่งเกี่ยวข้องกับประสิทธิภาพ คุณภาพของการให้บริการ และผลการดำเนินงานทางการเงิน เป้าหมายแห่งความสำเร็จ รวมทั้งตัวชี้วัดซึ่งได้ถูกออกแบบให้สนับสนุนการทำงานบนมิติทั้งสี่ของ Balanced Scorecard (BSC) โดยมีการกำหนดระยะเวลาในการประเมินและกำหนดน้ำหนักคะแนนที่แน่นอน สามารถวัดผลและนำผลที่ได้มาปรับปรุงการทำงานขององค์กรได้จริง

2. ทบทวนวรรณกรรมและบทความที่เกี่ยวข้อง

การวัดประสิทธิภาพขององค์กรมีวัตถุประสงค์เพื่อช่วยในการตัดสินใจในการกำหนดทิศทางบริหารจัดการ ช่วยให้พนักงานในองค์กรทราบถึงสิ่งที่ต้องปรับปรุงในการทำงาน ผลของการวัดประสิทธิภาพนอกจากช่วยให้องค์กรแก้ไขจุดบกพร่องได้ดีแล้ว ยังช่วยให้องค์กรเป็นที่น่าเชื่อถือในสายตาของบุคคลภายนอกและเป็นสิ่งจูงใจพนักงานให้ออกมาทำหน้าที่ของตนเองให้ดียิ่งขึ้น เครื่องมือที่ใช้วัดประสิทธิภาพขององค์กรมีหลากหลายในปัจจุบัน เช่น Value-based Management ที่ใช้แนวคิดการบริหารที่เน้นเพิ่มมูลค่าและผลประโยชน์ให้ผู้ถือหุ้นเป็นหลัก [1], Activity-based costing management เป็นแนวคิดการบริหารต้นทุนกิจกรรมใดๆ ในองค์กรให้คุ้มค่ากับการลงทุนมากที่สุด [2], Benchmarking ที่เน้นการเปรียบเทียบองค์กรของตนเองกับองค์กรที่ประกอบธุรกิจใกล้เคียงกันที่ประสบความสำเร็จเพื่อนำมาเป็นปรับปรุงการทำงาน [3] หรือ Balanced Scorecard ซึ่งเป็นเครื่องมือช่วยลัดช่องว่างระหว่างกลยุทธ์ขององค์กรและการดำเนินงานขององค์กร เป็นการวัดประสิทธิภาพที่ช่วยให้ผู้บริหารสามารถมองเห็นภาพรวมทาง

ธุรกิจและสามารถลงในรายละเอียดได้อย่างรวดเร็ว [4] การวัดประสิทธิภาพในงานวิจัยนี้เลือกใช้ BSC เป็นเครื่องมือใช้วัดและปรับปรุงยุทธศาสตร์ของแผนกให้บริการเนื่องจากส่วนการทำงานขององค์กรในภาพรวมมีการใช้เครื่องมือนี้อยู่ก่อนแล้ว ดังนั้นการใช้เครื่องมือเดียวกันก็จะทำให้สามารถนำเสนอผลการวัดประสิทธิภาพไปในทิศทางเดียวกัน นอกจากนี้ตัวเครื่องมือก็สามารถทำได้ง่ายและรวดเร็ว โดยใช้ข้อมูลพื้นฐานที่มีในองค์กรเพื่อกำหนดเป้าหมายความสำเร็จและตัวชี้วัด ตามมิติต่างๆ ประกอบด้วย 4 มิติ คือ 1) ด้านการเงิน (Financial) เช่น การเน้นลดต้นทุนให้กับองค์กร 2) ด้านความสัมพันธ์กับลูกค้า (Customer) เช่น การเน้นสร้างความพึงพอใจในการบริการ 3) ด้านกระบวนการภายใน (Internal Process) เช่น การพัฒนาขั้นตอนการทำงานให้มีประสิทธิภาพ และ 4) ด้านการเรียนรู้และการพัฒนา (Learning & Growth) เช่น การเน้นพัฒนาพนักงานให้มีความรอบรู้หรือมีความสามารถที่เป็นเลิศในหน้าที่ของตัวเอง เครื่องมือนี้ยังเอื้อให้พนักงานผู้มีส่วนร่วมในกระบวนการทำงานสามารถมีส่วนในการกำหนดเกณฑ์แห่งความสำเร็จได้ด้วยเพื่อให้เกิดความเต็มใจที่จะช่วยกันพัฒนาการทำงานให้ดีขึ้น และจากการศึกษางานวิจัยอื่นๆ พบว่ามีการนำ BSC และดัชนีตัวชี้วัดมาใช้ในการบริหารงานเพื่อความสำเร็จในหลายธุรกิจ [5,6,7] การใช้ BSC จะทำควบคู่กับการกำหนดเป้าหมาย (Target) การสร้างดัชนีตัวชี้วัด (KPI) และกำหนดเกณฑ์การให้คะแนนตัวชี้วัดดังจะแสดงในหัวข้อต่อไป

3. วิธีการดำเนินการวิจัย

งานวิจัยนี้ใช้วิธีการศึกษา (Case Study Methodology) [8] ขั้นตอนการให้บริการรับชำระค่าสินค้าและบริการของแผนกให้บริการ องค์กรด้านการสื่อสารและโทรคมนาคมแห่งหนึ่งในประเทศไทย โดยการศึกษาขั้นตอนการทำงานปัจจุบัน และใช้ข้อมูลจากระบบสารสนเทศตัวอย่าง ตั้งแต่ 1 มกราคม 2548 – 30 มิถุนายน 2549 เพื่อนำมากำหนดเป้าหมายและสร้างตัวชี้วัดที่สอดคล้องกับมิติต่างๆ ของ BSC และทดลองใช้ประเมินผลรอบไตรมาสที่สอง ปี 2549 เก็บข้อมูลด้วยการสัมภาษณ์ ศึกษาเอกสารและระบบสารสนเทศต่างๆ ที่เกี่ยวข้องกับการให้บริการรับชำระค่าสินค้าและบริการ เช่น ระบบ

บันทึกข้อร้องเรียนจากลูกค้า (Complaint Management System) ระบบบริหารจัดการคิว (Queue Management System) ระบบรับชำระค่าสินค้าและบริการ (Payment System) เป็นต้น

3.1 กำหนดเกณฑ์คะแนน KPI สำหรับตัวชี้วัด

การวัดประสิทธิภาพการให้บริการใช้วิธีการกำหนดเกณฑ์คะแนนของแต่ละตัวชี้วัดที่คำนวณได้จากสูตรต่างๆ จากนั้นนำค่าคะแนนที่ได้คูณกับน้ำหนัก KPI ของตัวชี้วัด เกณฑ์คะแนนที่ใช้มีค่าตั้งแต่ 1-5 อธิบายได้ดังนี้ 1-ต่ำกว่าเกณฑ์มาตรฐาน (Improved) 2-ต่ำกว่าเกณฑ์มาตรฐานแต่ยอมรับได้ (Accepted) 3-ตามเกณฑ์มาตรฐานขององค์กรที่ต้องดำเนินการให้ได้เป็นอย่างน้อย (Good) 4-มากกว่าเกณฑ์มาตรฐาน (Outstanding) 5-มากกว่าเกณฑ์มาตรฐานเป็นอย่างยิ่ง (Exceptional)

3.2 กำหนดเป้าหมายและเกณฑ์การให้คะแนน

ในการทำงานวิจัยนี้ ผู้วิจัยสัมภาษณ์ขั้นตอนการทำงานรวมทั้งศึกษาระบบสารสนเทศที่แผนกให้บริการใช้อยู่ในปัจจุบันพบว่า มีการเก็บข้อมูลที่หลากหลายและมีประโยชน์ เช่น ข้อมูลเงินเดือน ค่าใช้จ่ายเพื่อคำนวณต้นทุน ข้อมูลเวลาการให้บริการ ข้อมูลการร้องเรียนของลูกค้า เป็นต้น จึงนำข้อมูลที่มีมากำหนดเป้าหมายแห่งความสำเร็จได้ดังนี้

เป้าหมายตัวชี้วัดที่ 1 ควบคุมต้นทุนในการให้บริการลูกค้าต่อคนในเดือน ม.ค.-มิ.ย. ปี 2549 ให้ลดลง 20% โดยเปรียบเทียบกับต้นทุนในการให้บริการลูกค้าต่อคนในเดือน ม.ค.-มิ.ย. ปี 2548 ใช้เกณฑ์คะแนนดังนี้ 1 คะแนน ถ้าต้นทุนลดลงไม่เกิน 5%, 2 คะแนน ถ้าต้นทุนลดลงตั้งแต่ 6-10%, 3 คะแนน ถ้าต้นทุนลดลงตั้งแต่ 11-15%, 4 คะแนน ถ้าต้นทุนลดลงตั้งแต่ 16-19% และ 5 คะแนน ถ้าต้นทุนลดลงตั้งแต่ 20% ขึ้นไป (สมการ (1))

เป้าหมายตัวชี้วัดที่ 2 ควบคุมปริมาณปัญหาที่ลูกค้าร้องเรียนในไตรมาส 2 ปี 2549 ให้ลดลง 20% โดยเปรียบเทียบกับปริมาณปัญหาที่ลูกค้าร้องเรียนในไตรมาส 1 ปี 2549 ใช้เกณฑ์คะแนนดังนี้ 1 คะแนน ถ้าปัญหาลดลงไม่เกิน 5%, 2 คะแนน ถ้าปัญหาลดลงตั้งแต่ 6-10%, 3 คะแนน ถ้าปัญหาลดลงตั้งแต่ 11-15%, 4 คะแนน ถ้าปัญหาลดลงตั้งแต่ 16-19% และ 5 คะแนน ถ้าปัญหาลดลงตั้งแต่ 20% ขึ้นไป (สมการ (2))

เป้าหมายตัวชี้วัดที่ 3 ควบคุมระยะเวลาการให้บริการลูกค้าเฉลี่ยต่อรายการ ในไตรมาส 2 ปี 2549 ให้มีการเพิ่มขึ้นหรือลดลงไม่เกิน 0.15 นาที เปรียบเทียบกับระยะเวลาการให้บริการลูกค้าเฉลี่ยต่อรายการในไตรมาส 1 ปี 2549 ใช้เกณฑ์คะแนนดังนี้ 1 คะแนน ถ้าเวลาเฉลี่ยมากกว่า 1 นาที, 2 คะแนน ถ้าเวลาเฉลี่ยตั้งแต่ 0.46-1.00 นาที, 3 คะแนน ถ้าเวลาเฉลี่ยตั้งแต่ 0.31-0.45 นาที, 4 คะแนน ถ้าเวลาเฉลี่ยตั้งแต่ 0.16-0.30 นาที และ 5 คะแนน ถ้าเวลาเฉลี่ยไม่เกิน 0.15 นาที (สมการ (3))

เป้าหมายตัวชี้วัดที่ 4 ควบคุมจำนวน Transaction ที่ประมวลผลเกิน 5 วินาทีในไตรมาส 2 ปี 2549 ให้ลดลง 10% โดยเปรียบเทียบกับจำนวน Transaction ที่ประมวลผลเกิน 5 วินาทีในไตรมาส 1 ปี 2549 ใช้เกณฑ์คะแนนดังนี้ 1 คะแนน ถ้าประมวลผลลดลงไม่เกิน 0%, 2 คะแนน ถ้าประมวลผลลดลงตั้งแต่ 1-3%, 3 คะแนน ถ้าประมวลผลลดลงตั้งแต่ 4-6%, 4 คะแนน ถ้าประมวลผลลดลงตั้งแต่ 7-9% และ 5 คะแนน ถ้าประมวลผลลดลงตั้งแต่ 10% ขึ้นไป (สมการ (4))

เป้าหมายตัวชี้วัดที่ 5 ควบคุมจำนวนพนักงานที่ถูกร้องเรียนในไตรมาส 2 ปี 2549 ให้ลดลง 80% โดยเปรียบเทียบกับจำนวนพนักงานที่ถูกร้องเรียนในไตรมาส 1 ปี 2549 ใช้เกณฑ์คะแนนดังนี้ 1 คะแนน ถ้าจำนวนพนักงานถูกร้องเรียนลดลงไม่เกิน 0%, 2 คะแนน ถ้าจำนวนพนักงานถูกร้องเรียนลดลงตั้งแต่ 1-30%, 3 คะแนน ถ้าจำนวนพนักงานถูกร้องเรียนลดลงตั้งแต่ 31-60%, 4 คะแนน ถ้าจำนวนพนักงานถูกร้องเรียนลดลงตั้งแต่ 61-79% และ 5 คะแนน ถ้าจำนวนพนักงานถูกร้องเรียนลดลงตั้งแต่ 80% ขึ้นไป (สมการ (5))

3.3 สูตรตัวชี้วัดที่ใช้เพื่อวัดประสิทธิภาพการให้บริการ

ตัวชี้วัดที่ 1 : ต้นทุนในการให้บริการลูกค้าต่อคน (บาท)

หมายถึง ต้นทุนที่ทางองค์กรให้บริการลูกค้าต่อคน คำนวณโดยใช้ข้อมูลเงินเดือนของพนักงาน จำนวนรายการที่ดำเนินการและค่าใช้จ่ายคงที่

$$KPI1 = \left(\frac{\sum_{m=jan}^n \frac{a \times b + d}{c}}{n} \right) \quad (1)$$

เมื่อ a เงินเดือนเฉลี่ยของพนักงาน
b จำนวนพนักงาน
c จำนวนรายการของเดือน m
d ค่าใช้จ่ายคงที่ต่อเดือน

m เดือนที่ต้องการเริ่มคิดต้นทุน
n จำนวนเดือนที่ต้องการเปรียบเทียบ

ตัวชี้วัดที่ 2 : ปริมาณปัญหาที่ลูกค้าร้องเรียน (จำนวนครั้ง)
หมายถึง ปริมาณปัญหาที่ลูกค้ามีการร้องเรียนเข้ามา ซึ่งปัญหาเหล่านี้ได้รับการบันทึกไว้ในระบบเรียบร้อยแล้ว

$$KPI2 = \left(\frac{e-f}{e}\right) \times 100 \quad (2)$$

เมื่อ e จำนวนครั้งของปัญหาที่ถูกร้องเรียนไตรมาส 1
f จำนวนครั้งของปัญหาที่ถูกร้องเรียนไตรมาส 2

ตัวชี้วัดที่ 3 : ระยะเวลาที่ใช้ในการให้บริการลูกค้า เฉลี่ย/คน/ไตรมาส (นาที)

หมายถึง ระยะเวลาที่ทางองค์กรให้บริการลูกค้าต่อคนจับเวลาด้วยโปรแกรมชื่อว่า Q-Matic เวลาที่ทำการจับและบันทึกเพื่อนำมาหาผลรวมคือ Client Time + Server Time + Network Time

$$KPI3 = |g - h| \quad (3)$$

เมื่อ g ระยะเวลาที่ใช้ในการให้บริการลูกค้าเฉลี่ยรายการ/ไตรมาส 2
h ระยะเวลาที่ใช้ในการให้บริการลูกค้าเฉลี่ยรายการ/ไตรมาส 1

ตัวชี้วัดที่ 4 : จำนวน Transaction ที่ประมวลผลเกิน 5 วินาที (Transaction)

หมายถึง การนับจำนวน Transaction ที่ประมวลผลเกิน 5 วินาที ด้วยโปรแกรมชื่อว่า End2End โดยนับเวลาที่ Title Bar ของ Internet Explorer เปลี่ยนจากหน้าหนึ่งไปอีกหน้าหนึ่ง

$$KPI4 = \left(\frac{k-l}{k}\right) \times 100 \quad (4)$$

เมื่อ k จำนวน Transaction ที่เกิน 5 วินาทีในไตรมาส 1
l จำนวน Transaction ที่เกิน 5 วินาทีในไตรมาส 2

ตัวชี้วัดที่ 5 : จำนวนพนักงานที่ถูกร้องเรียน (คน)

หมายถึง จำนวนพนักงานที่ถูกร้องเรียนโดยลูกค้า และทางสาขาหรือหัวหน้า หรือผู้มีอำนาจได้พิจารณาแล้ว เห็นสมควรให้บันทึกว่าถูกร้องเรียนจริง เช่น พนักงานดำเนินการเรื่องเงินของลูกค้านี้ไม่ถูกต้อง

$$KPI5 = \left(\frac{p-q}{p}\right) \times 100 \quad (5)$$

เมื่อ p จำนวนพนักงานที่ถูกร้องเรียนในไตรมาส 1
q จำนวนพนักงานที่ถูกร้องเรียนในไตรมาส 2

4. ผลการดำเนินงานและการวิเคราะห์ผล

4.1 มิติ BSC กับเป้าหมายตัวชี้วัด

จากข้อมูลระบบสารสนเทศที่มีอยู่ภายในองค์กรของแผนกให้บริการรับชำระค่าสินค้าและบริการ เราสามารถระบุเป้าหมายตัวชี้วัดที่สอดคล้องกับมิติของ BSC ได้ตามตารางที่ 1 โดยในแต่ละตัวชี้วัดมีการกำหนดน้ำหนักคะแนน KPI เป็น 20% ทั้งนี้เป็นเพราะองค์กรได้ให้ความสำคัญกับตัวชี้วัดในแต่ละตัวเท่าๆกัน

ตารางที่ 1 มิติ BSC กับเป้าหมายตัวชี้วัด

BSC	เป้าหมายตัวชี้วัด
มิติด้านการเงิน	1. เพื่อควบคุมต้นทุนในการให้บริการลูกค้า
มิติด้านผู้รับบริการ	2. เพื่อควบคุมปริมาณปัญหาที่ลูกค้าร้องเรียน
มิติด้านกระบวนการ	3. เพื่อควบคุมระยะเวลาการให้บริการลูกค้าเฉลี่ยต่อรายการ 4. เพื่อควบคุมจำนวน Transaction ที่ประมวลผลเกิน 5 วินาที
มิติด้านการเรียนรู้	5. เพื่อควบคุมจำนวนพนักงานที่ถูกร้องเรียน

4.2 ผลการวัดประสิทธิภาพแต่ละตัวชี้วัด

ในขั้นตอนนี้ผลที่ได้เกิดจากการนำข้อมูลสารสนเทศที่มีการเก็บบันทึกไว้จริงมาคำนวณด้วยสูตรตัวชี้วัดที่ผู้วิจัยได้นำเสนอในหัวข้อ 3.3 และให้คะแนนตามเกณฑ์ที่ตั้งไว้ในหัวข้อ 3.2 ได้ผลการวัดประสิทธิภาพดังนี้

ตัวชี้วัดที่ 1 ต้นทุนการให้บริการในเดือน ม.ค.- มิ.ย. 2548 และ เดือน ม.ค.- มิ.ย. 2549 เป็น 21.85 และ 16.74 มีค่าลดลง 23.38% ได้คะแนนประเมิน 5 คะแนนซึ่งถือว่าอยู่ในระดับที่ดีมาก ทั้งนี้เนื่องจากในระยะที่คำนวณต้นทุน องค์กรไม่ได้ใช้เงินลงทุนเพิ่มขึ้นเพื่อดำเนินการใดๆ นอกจากนี้ปริมาณลูกค้าที่มาเข้ารับบริการก็มีมากขึ้นในขณะที่พนักงานผู้ให้บริการยังมีเท่าเดิม (แต่ยังคงสามารถรองรับลูกค้าได้) เพื่อคงต้นทุนให้ลดลงได้อีกพนักงานควรให้ความร่วมมือในทุกด้านเพื่อดำเนินการ

ตัวชี้วัดที่ 2 ปริมาณปัญหาที่ลูกค้าร้องเรียนใน Q1 มีจำนวน 1,752 และใน Q2 มีจำนวน 2,157 เพิ่มขึ้น 23.12% ได้คะแนนประเมิน 1 คะแนน ดังตารางที่ 2 จะพบว่าค่าที่ได้ต่ำกว่าเกณฑ์ที่

ยอมรับได้ สาเหตุอาจจะเกิดจากปริมาณลูกค้าที่เพิ่มขึ้นอย่างต่อเนื่องและยังมีการเปิดให้บริการใหม่ๆ ที่พนักงานยังไม่ได้เรียนรู้ ส่งผลให้ไม่สามารถอธิบายรายละเอียด หรือแจ้งรายละเอียดไม่ถูกต้องแก่ลูกค้า แนวทางในการปรับปรุง คือ ต้องพยายามลดการร้องเรียนให้มากที่สุด ปลูกฝังพนักงานให้มีใจบริการ และเมื่อมีบริการใหม่ๆเกิดขึ้นควรให้ความรู้กับพนักงานผู้ให้บริการก่อนเปิดบริการจริงให้ลูกค้า

ตารางที่ 2 ผลตัวชี้วัดที่ 2

ประเภทปัญหา	จำนวน(ครั้ง)	
	Q1	Q2
ไม่ใช้ KM /PROCESS	247	241
ข้อมูลใน KM /PROCESS ไม่ชัดเจน	26	30
ไม่มีข้อมูลที่ KM	4	8
SYSTEM ERROR	25	29
พนักงานทำผิดระเบียบวินัย	138	148
พนักงานขาดความรอบคอบในการปฏิบัติงาน	1,251	1,604
พนักงานไม่ยืดหยุ่นต่อสถานการณ์	53	73
พนักงาน OJT ไม่ทราบข้อมูล	8	24
รวม	1,752	2,157

ตัวชี้วัดที่ 3 ระยะเวลาการให้บริการลูกค้าเฉลี่ยต่อรายการมีผลต่างระหว่างไตรมาส 34 วินาทีโดยใน Q1 มีระยะเวลาเฉลี่ย 5.49 นาที และใน Q2 มีระยะเวลา 6.23 นาที ได้คะแนนประเมิน 3 คะแนนถือว่าอยู่ในเกณฑ์มาตรฐานของแผนก กล่าวคือสามารถรักษาเวลาของการให้บริการลูกค้าต่อคนได้ในช่วง 5-7 นาทีโดยเฉลี่ย แต่เพื่อการปรับปรุงบริการให้เร็วขึ้น พนักงานควรมีคำแนะนำหรือเอกสารแนะนำบริการใหม่ๆ สำหรับการเตรียมตัวก่อนเข้ารับบริการให้ลูกค้ารับทราบในระหว่างรอคอยเพื่อช่วยลดระยะเวลาการดำเนินการ การนำระบบไอทีเข้ามาใช้ เช่น การให้ลูกค้าบริการตัวเอง ในเบื้องต้นก็ถือเป็นอีกวิธีหนึ่งที่จะช่วยให้ลดเวลาการรอคอยได้มากขึ้น

ตัวชี้วัดที่ 4 จำนวน Transaction ที่ประมวลผลเกิน 5 วินาที ใน Q1 จำนวน 371 รายการ และใน Q2 จำนวน 309 รายการ

ถือว่าลดลง 16.71 % ได้คะแนนประเมิน 5 คะแนนอยู่ในระดับดีมาก ทั้งนี้เนื่องจากระบบสารสนเทศที่มีความทันสมัยมากขึ้น แต่เพื่อการพัฒนาอย่างต่อเนื่องควรตรวจสอบประสิทธิภาพของระบบให้บริการลูกค้าในทุกๆระบบ และเน้นย้ำการติดตาม (Monitor) ในช่วงที่มีการใช้บริการในปริมาณมากอย่างสม่ำเสมอ

ตัวชี้วัดที่ 5 จำนวนพนักงานที่ถูกร้องเรียนลดลง 1 คนจาก Q1 ซึ่งมีการแจ้งร้องเรียน 3 คนคิดเป็นลดลง 33.33% ได้คะแนนประเมิน 3 คะแนนอยู่ในระดับมาตรฐานที่ตั้งไว้ ทั้งนี้เพราะมีการนำระบบสารสนเทศมาใช้มากขึ้นทำให้พนักงานสามารถตอบคำถามและแก้ไขปัญหาได้ดี แต่อย่างไรก็ตามการป้องกันไม่ให้เกิดข้อร้องเรียนจะดีที่สุดกับองค์กรในด้านภาพลักษณ์การให้บริการ องค์กรอาจจะมีบทลงโทษสำหรับพนักงานที่ถูกร้องเรียน และมีการสร้างขวัญและกำลังใจสำหรับพนักงานที่ทำหน้าที่บริการที่ดีด้วย

4.3 ผลการวัด KPI ของแผนกให้บริการรับชำระค่าสินค้าและบริการ

การวัดประสิทธิภาพในภาพรวมของแผนกให้บริการรับชำระค่าสินค้าและบริการได้ผลลัพธ์เท่ากับ 3.4 ดังตารางที่ 3 ถือว่าอยู่ในเกณฑ์มาตรฐานขององค์กร

ตารางที่ 3 ผลการวัดประสิทธิภาพของแผนกให้บริการ

ตัวชี้วัดที่	น้ำหนัก (W)	คะแนนประเมิน (S)	คะแนนจริง (W x S)
1	20	5	100
2	20	1	20
3	20	3	60
4	20	5	100
5	20	3	60
รวม	100		340
KPI = ผลรวมคะแนนจริง/ผลรวมน้ำหนัก			3.4

5. บทสรุปและข้อเสนอแนะ

งานวิจัยการออกแบบตัวชี้วัดประสิทธิภาพการให้บริการรับชำระค่าสินค้าและบริการเป็นการนำเป้าหมายและตัวชี้วัดมาวัดประสิทธิภาพการทำงานของแผนกให้บริการโดยใช้เครื่องมือ BSC และ KPI ในการวัด มีการกำหนดตัวชี้วัดและสูตรเพื่อคำนวณค่าตัวชี้วัดต่างๆ ผลการนำไปใช้ที่นำเสนอในหัวข้อ 4 ผลการดำเนินงาน ภาพรวมของแผนกได้คะแนนรวมทั้ง

3.4 ซึ่งถือว่าอยู่ในเกณฑ์มาตรฐานที่องค์กรมองว่าทำงานได้ดีมีประสิทธิภาพ นอกจากนี้ผลคะแนนในแต่ละตัวชี้วัดยังสามารถบอกได้ว่า แผนกมีการทำงานที่ดีอยู่แล้วในบางด้าน และยังมีประเด็นอะไรที่ต้องปรับปรุงให้ดียิ่งขึ้น แม้ข้อมูลที่น่าสนใจในการศึกษาครั้งนี้ จะเป็นข้อมูลจากการดำเนินงานในหลายปีก่อน แต่ผลการวิจัยพอสรุปได้ว่า แนวคิดในการกำหนดเป้าหมายและตัวชี้วัดซึ่งได้จากการนำข้อมูลที่เก็บไว้จากการปฏิบัติงานจริงในแต่ละขั้นตอนมาใช้ มีประโยชน์และสามารถนำไปประยุกต์ใช้ได้จริงกับองค์กร

จะเห็นได้ว่า การที่องค์กรใดมีการวัดประสิทธิภาพการทำงานขององค์กร จะทำให้ได้รับประโยชน์จากผลที่วัดประสิทธิภาพได้และนำมาซึ่งการปรับปรุงกลยุทธ์การดำเนินงานของกิจการ รวมทั้งช่วยในการกำหนดยุทธศาสตร์ขององค์กรที่จะนำมาใช้ในอนาคตเพื่อแข่งขันกับคู่แข่งได้ นอกจากนี้ยังเป็นตัวกระตุ้นให้ผู้ร่วมงานช่วยกันพัฒนาการทำงานให้มีประสิทธิภาพอย่างต่อเนื่องเนื่องจากทราบแล้วว่า องค์กรยังต้องการการพัฒนาในส่วนใดเพิ่มเติมบ้าง ตัววัดประสิทธิภาพในแต่ละองค์กรย่อมมีความแตกต่างกันเล็กน้อยแล้วแต่สภาพแวดล้อมการทำงานในองค์กรนั้นๆ

การวัดประสิทธิภาพในองค์กรใดๆ จะประสบผลสำเร็จได้ ต้องได้รับความร่วมมือทุกภาคส่วน เช่น ผู้บริหารสูงสุดขององค์กร จะต้องเห็นความสำคัญของการวัดประสิทธิภาพของหน่วยงาน และต้องมีความเชื่อมั่นว่า การวัดประสิทธิภาพจะนำมาซึ่งการพัฒนาองค์กรที่ดีขึ้นในอนาคต หัวหน้าในแผนกที่ได้รับมอบหมายจากผู้บริหารให้ดูแลเรื่องการวัดประสิทธิภาพของหน่วยงาน จะต้องรับนโยบายมาจากผู้บริหาร และต้องเป็นผู้ร่วมดำเนินการในการวัดประสิทธิภาพของแผนกของตนอย่างจริงจังและอธิบายให้พนักงานเข้าใจว่า การวัดประสิทธิภาพของแผนกมีผลช่วยเพิ่มกำไรขององค์กรให้มากขึ้น ซึ่งเกี่ยวข้องกับสวัสดิภาพการทำงานของพวกเขาโดยตรง ส่วนผู้กำหนดตัวชี้วัดเพื่อใช้ในการวัดประสิทธิภาพ จะต้องมีความเข้าใจในการทำงานของแผนกนั้นๆอย่างแท้จริง จึงจะสามารถกำหนดตัวชี้วัด ค่าเป้าหมาย และแปลความหมายของการวัดประสิทธิภาพได้อย่างถูกต้องมากที่สุด และเพื่อให้การวิเคราะห์ผลเป็นไปอย่างถูกต้อง องค์กรควรกำหนดระยะเวลาการวัดผลด้วยว่าจะวัดผลทุกๆระยะกี่วัน กี่เดือน หรือ กี่ปี เพื่อจะได้

ผลตัวชี้วัดที่ได้ไปปรับใช้ในการดำเนินงานขององค์กรต่อไป และควรมีระบบสารสนเทศที่จะช่วยในการกำหนดค่าเป้าหมาย การกำหนดตัวชี้วัด และการวัดประสิทธิภาพต่างๆขององค์กร เพื่อช่วยลดระยะเวลาในการเก็บรวบรวมข้อมูล และการวิเคราะห์ผล

งานวิจัยนี้ บรรลุวัตถุประสงค์ในการจัดทำเป้าหมาย ตัวชี้วัด และสูตรการคำนวณ สามารถนำไปใช้ได้จริง งานชิ้นต่อไปที่ผู้วิจัยวางแผนจะดำเนินการ คือ การพัฒนาระบบสารสนเทศเพื่อรองรับการนำข้อมูลเข้าจากระบบต่างๆ เพื่อแสดงผลตัวชี้วัด โดยอัตโนมัติให้องค์กรนำไปใช้ต่อไป

เอกสารอ้างอิง

- [1] Ittner, Christopher D., and David F. Larcker, "Assessing empirical research in managerial accounting: a value-based management perspective," *Journal of accounting and economics*, vol. 32, no. 1-3, pp. 349-410, 2001.
- [2] Armstrong, Peter, "The costs of activity-based management," *Accounting, Organizations and Society*, vol. 27, no. 1, pp. 99-120, 2002.
- [3] Zairi, Mohamed, and Rob Hutton., "Benchmarking: a process-driven tool for quality improvement," *The TQM Magazine*, vol. 7, no. 3, pp. 35-40, 1995.
- [4] Kaplan, Robert S., and David P. Norton, *Balanced scorecard.*: Gabler, 2007.
- [5] Davis, Stan, and Tom Albright, "An investigation of the effect of balanced scorecard implementation on financial performance," *Management Accounting Research*, vol. 15, no. 2, pp. 135-153, 2004.
- [6] Inamdar, Noorein, Robert S. Kaplan, and Marvin Bower, "Applying the balanced scorecard in healthcare provider organizations," *Journal of healthcare management/American College of Healthcare Executives*, vol. 47, no. 3, p. 179, 2002.
- [7] Ahn, Heinz, "Applying the balanced scorecard concept: an experience report," *Long range planning*, vol. 34, no. 4, pp. 441-461, 2001.
- [8] Baxter, Pamela, and Susan Jack, "Qualitative case study methodology: Study design and implementation for novice researchers," *The qualitative report*, vol. 13, no. 4, pp. 544-559, 2008.

การประยุกต์ใช้สื่อ eDLTV เพื่อการเรียนรู้สู่การพัฒนาแอปพลิเคชันเกมเพื่อการเรียนรู้บนแท็บเล็ต

The Application of eDLTV Media to Develop the Games Application for Learning on Tablet

พีรชัย ชินพร (peerachai chinporn)¹ และ วรปภา อารีราษฎร์ (Worapapha Arreerard)²

¹สาขาคอมพิวเตอร์ศึกษา คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏมหาสารคาม

²คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏมหาสารคาม

¹ Koopee27@hotmail.com, ²dr.Worapapha@hotmail.com

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ 1) พัฒนาหลักสูตรอบรม การประยุกต์ใช้สื่ออีดีแอลทีวีเพื่อการเรียนรู้สู่การพัฒนา แอปพลิเคชันเกมเพื่อการเรียนรู้บนแท็บเล็ต 2) ศึกษาความรู้ และความเข้าใจของครูที่เข้ารับการอบรมตามหลักสูตรอบรมที่ พัฒนาขึ้น 3) ศึกษาคุณภาพของแอปพลิเคชันเกมเพื่อการเรียนรู้ บนแท็บเล็ตของที่ครูพัฒนาขึ้น 4) ศึกษาความพึงพอใจของครู ที่เข้ารับการอบรม กลุ่มเป้าหมาย คือครูผู้สอนชั้นประถมศึกษา ปีที่ 1 สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษามหาสารคาม เขต 1 จำนวน 30 คน เครื่องมือที่ใช้ในการวิจัย ได้แก่ หลักสูตร แบบประเมินหลักสูตรการฝึกอบรมเชิงปฏิบัติการ แบบทดสอบวัดผลการเรียนรู้ของผู้เข้ารับการอบรมเชิงปฏิบัติการ แบบประเมินแอปพลิเคชันเกมเพื่อการเรียนรู้บนแท็บเล็ต และแบบประเมินความพึงพอใจของผู้เข้ารับการอบรม ที่มีต่อหลักสูตรอบรม สถิติที่ใช้วิเคราะห์ข้อมูล คือ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และสถิติ *t-test* แบบ (Dependent Sample)

ผลการวิจัยพบว่า 1) หลักสูตรอบรมการประยุกต์ใช้สื่ออีดีแอลทีวีเพื่อการเรียนรู้สู่การพัฒนาแอปพลิเคชันเกมเพื่อการเรียนรู้บนแท็บเล็ต ประกอบด้วยกระบวนการอบรม 5 ขั้นตอน ขั้นที่ 1 เรียนรู้จากการรับฟังการบรรยาย ขั้นที่ 2 เรียนรู้จากการสาธิตการใช้สื่ออีดีแอลทีวี และสาธิตการสร้างแอปพลิเคชันเกมส์ ขั้นที่ 3 เรียนรู้การสร้างแอปพลิเคชันเกมส์ จาก การปฏิบัติการ ขั้นที่ 4 เรียนรู้จากการให้คำแนะนำของวิทยากร และผู้ช่วยวิทยากร ขั้นที่ 5 ปรับปรุงแก้ไขแอปพลิเคชันเกมส์ ผู้เชี่ยวชาญมีความคิดเห็นโดยรวมในระดับเหมาะสมมากที่สุด ($\bar{X}=4.5$, $S.D.=0.49$) 2) ผู้เข้ารับการอบรมมีคะแนนความรู้ และความเข้าใจหลังการอบรมสูงกว่าก่อนการอบรม อย่างมี

นัยสำคัญทางสถิติที่ระดับ .05 3) คุณภาพของแอปพลิเคชันที่ครูพัฒนาขึ้น มีคุณภาพอยู่ในระดับมากถึงมากที่สุด และ 4) ครูที่เข้ารับการอบรมมีความพึงพอใจต่อหลักสูตรอบรมอยู่ในระดับพึงพอใจระดับมากที่สุด ($\bar{X}=4.67$, $S.D.=0.47$)

คำสำคัญ : สื่อ eDLTV แอปพลิเคชัน แท็บเล็ต เกมเพื่อการเรียนรู้

Abstract

This research aims to 1) Developing the training syllabus of eDLTV application to games developed application for learning on the tablet 2) to study the learning and understanding of the participated teachers in this training syllabus 3) to study the quality of performance of games developed application for learning of the participated teachers 4) to study the satisfaction of the participated teachers in this program. 30 teachers who teach in primary 1 of the office of Maha Sarakham Primary Education Service Area 1 were selected as the sample group. The research tools were course syllabus, the assessment of workshop, the test to assess the learning perception of the participated teachers, the assessment for games developed application of participated teachers, and the satisfaction questionnaire for the participated teacher for the syllabus. The statistics used in data analysis were Average, Standard Deviation and *t-test* (Dependent Sample).

The results of this research were found that 1) The training syllabus of eDLTV application to games developed application for learning on the tablet consisted of 5 stages: 1 learned from lectures 2 Learned from the demonstration of eDLTV media and demonstrates how to create games application 3 learned how to create games

application by workshop 4 Learned from the instruction of lecturer and assistant lecturer 5 improved the games application. The specialists had estimated overview at the highest appreciated level ($\bar{X}=4.5$, $S.D.=0.49$) 2) The participated teachers were acknowledged and understand after the training syllabus had Average mean score at 0.05 levels statistically significant. 3) The Quality of developed application had high quality level to highest level. and 4) The participated teachers have satisfaction on the developed syllabus at the highest score ($\bar{X} =4.67$, $S.D.=0.47$)

Keywords : eDLTV Media, Applications, Tablet, Games for Learning

1. บทนำ

นโยบายภาครัฐด้านการจัดการศึกษาของรัฐบาลด้านการพัฒนาการใช้เทคโนโลยีสารสนเทศเพื่อการศึกษาให้ทัดเทียมกับนานาชาตินั้น เป็นนโยบายที่มีความสำคัญยิ่ง ได้กำหนดแนวนโยบายเพื่อเร่งพัฒนาการใช้เทคโนโลยีสารสนเทศเพื่อการศึกษาให้เป็นเครื่องมือยกระดับคุณภาพและกระจายโอกาสทางการศึกษาให้มีระบบการเรียนแบบอิเล็กทรอนิกส์แห่งชาติเป็นกลไกในการปรับเปลี่ยนกระบวนการเรียนรู้ที่เอื้อให้เกิดการเรียนรู้ตลอดชีพ ส่งเสริมให้นักเรียนทุกระดับชั้นใช้อุปกรณ์คอมพิวเตอร์แท็บเล็ตเพื่อการศึกษา โดยได้จัดสรรเครื่องคอมพิวเตอร์แท็บเล็ตให้กับนักเรียนชั้นประถมศึกษาปีที่ 1 ในการปีการศึกษา 2555 [1]

สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน หรือ สพฐ. ได้จัดทำโครงการจัดการประกวดและอบรมเชิงปฏิบัติการผลิตสื่อการเรียนการสอนเพื่อใช้กับเครื่องแท็บเล็ต โดยจัดเตรียมเครื่องมือสร้างสื่อ และอบรมให้กับผู้เข้าประกวดโดยแบ่งเป็น 2 ประเภท ได้แก่ ครูและบุคลากรทางการศึกษา จัดอบรม 4 วัน 3 คืน ครั้งละ 200 คน ใน 5 จังหวัดทั่วประเทศ และจัดอบรมสำหรับบุคคลธรรมดาและนิติบุคคล จัดอบรมเชิงสัมมนา 5 ชั่วโมง และการอบรมผ่านระบบเครือข่าย แบบ e-Training [2]

คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏมหาสารคาม ได้ดำเนินการร่วมมือกับเครือข่ายทั้งหน่วยงานภาครัฐและเอกชน จากผลการดำเนินงานโครงการส่งเสริมการใช้สื่ออีดีแอลทีวี (eDLTV : Electronic Distance Learning Via Television) ซึ่งเป็นสื่อภายใต้โครงการจัดทำเนื้อหาบนระบบอี

ดีแอลทีวี (eDL-Square) ภายใต้โครงการเทคโนโลยีสารสนเทศตามพระราชดำริ สมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี เพื่อให้ครูได้ใช้เป็นสื่อเพื่อการเรียนการสอน จากการสำรวจโรงเรียนและความต้องการของครูในการใช้สื่ออีดีแอลทีวีเพื่อการเรียนการสอนพบว่า ครูและบุคลากรทางการศึกษามีความสนใจในการนำสื่ออีดีแอลทีวี ไปใช้ในการจัดการเรียนรู้ร้อยละ 92.44 โดยนำไปเป็นแหล่งสืบค้น/แหล่งเรียนรู้ ร้อยละ 88.02 และเป็นสื่อเสริม ร้อยละ 82.23 และครูและบุคลากรทางการศึกษามีความสนใจในการพัฒนาต่อยอดการจัดการเรียนรู้ด้วยสื่ออิเล็กทรอนิกส์ นิกส์ ร้อยละ 86.23 เนื่องจากสื่ออีดีแอลทีวีเป็นสื่อที่มีเนื้อหาสาระที่ครอบคลุมการจัดการเรียนการสอนในระดับประถม ศึกษา [3]

การส่งเสริมการพัฒนาครูในการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนการสอนตามนโยบายที่ได้ดำเนินการผ่านมาเป็น การดำเนินการนำร่องตามนโยบายของ สพฐ. ยังไม่ครอบคลุมครูและบุคลากรทางการศึกษาทั่วประเทศ และจากการที่มหาวิทยาลัยราชภัฏมหาสารคาม ได้ดำเนินการขยายผลการอบรม อบรมครูร่วมกับสำนักงานเขตพื้นที่การศึกษาในเขตพื้นที่บริการ จากการใช้บริการอบรมเดือน พฤษภาคม 2555 ถึง เดือน พฤษภาคม 2556 จำนวน 8 รุ่น มีผู้เข้ารับการอบรมรวมทั้งหมด 493 คน ก็ยังไม่ครอบคลุมครูในเขตพื้นที่บริการ ส่งผลให้ครูผู้สอนสังกัดสำนักงานเขตพื้นที่การศึกษา มหาสารคามเขต 1 มีความสนใจเข้ารับการอบรมอย่างต่อเนื่อง อีกทั้งนักเรียนที่อยู่ในระดับประถมศึกษา การเรียนรู้ด้วยเกมเพื่อการศึกษา เป็นสื่อที่เหมาะสมและสอดคล้องกับความต้องการและวัยของผู้เรียน นักเรียนจะมีความกระตือรือร้นในการเรียน สร้างความสนใจและทำให้ผู้เรียน มีความตั้งใจในการเรียนมากยิ่งขึ้น [4]

จากหลักการและเหตุผลดังกล่าวข้างต้น ผู้วิจัยได้เห็นความสำคัญและความจำเป็นในการอบรมครูพัฒนาแอปพลิเคชันเกมเพื่อการเรียนรู้บนแท็บเล็ต โดยการศึกษาเนื้อหาสาระการเรียนรู้จากสื่ออีดีแอลทีวี นำมาสู่ออกแบบเกมเพื่อการศึกษา เพื่อให้ครูสามารถสร้างแอปพลิเคชันเกมเป็นสื่อในการเรียนการสอน ให้กับนักเรียนชั้นประถมศึกษาได้อย่างมีคุณภาพ ส่งผลให้นักเรียนมีสื่อที่เป็นแอปพลิเคชันแบบเกมที่สามารถใช้ในการเรียนรู้บนแท็บเล็ตเพิ่มมากขึ้น และครูผู้สอนที่เข้ารับการอบรมมีความพึงพอใจในกระบวนการอบรมการพัฒนาแอปพลิเคชันสู่แท็บเล็ต และงานวิจัยนี้ยังส่งผลให้มหาวิทยาลัย

ราชภัฏมหาสารคาม ได้ให้บริการชุมชนที่สอดคล้องกับนโยบายของ สพฐ.ที่มหาวิทยาลัยราชภัฏมหาสารคามเป็นหน่วยงานหลัก ในการส่งเสริมการพัฒนาศักยภาพครูและบุคลากรทางการศึกษา ในเขตพื้นที่บริการอย่างต่อเนื่อง

วัตถุประสงค์

1. เพื่อพัฒนาหลักสูตรอบรมการประยุกต์ใช้สื่ออีดีแอลเพื่อการเรียนรู้สู่การพัฒนาแอปพลิเคชันเกมเพื่อการเรียนรู้บนแท็บเล็ต
2. เพื่อศึกษาความรู้และความเข้าใจของครูที่เข้ารับการอบรมตามหลักสูตรอบรมที่พัฒนาขึ้น
3. เพื่อศึกษาคุณภาพของแอปพลิเคชันเกมเพื่อการเรียนรู้บนแท็บเล็ตของที่ครูพัฒนาขึ้น
4. ศึกษาความพึงพอใจของครูที่เข้ารับการอบรม

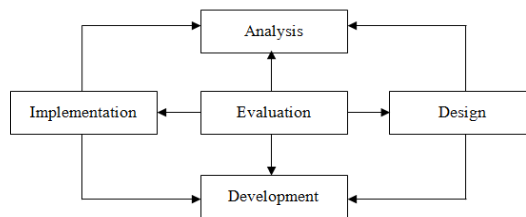
2. วรรณกรรมที่เกี่ยวข้อง

2.1 สื่ออีดีแอลทีวี

สื่ออีดีแอลทีวี หรือ eDLTV (Electronic Distance Learning Television) เป็นสื่อที่พัฒนาโดยโครงการจัดทำเนื้อหา ระบบ e-Learning ของการศึกษาทางไกลผ่านดาวเทียม เฉลิมพระเกียรติเนื่องในโอกาสมหามงคลเฉลิมพระชนมพรรษา 80 พรรษา 5 ธันวาคม 2550 เป็นโครงการความร่วมมือของ มูลนิธิการศึกษาทางไกลผ่านดาวเทียม ร่วมกับโครงการเทคโนโลยีสารสนเทศตามพระราชดำริสมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี โดยการนำเนื้อหาของการศึกษาทางไกลผ่านดาวเทียม (DLTV) ที่ออกอากาศทางสถานีวิทยุและโทรทัศน์การศึกษาทางไกลผ่านดาวเทียม จากโรงเรียนวังไกลกังวล จังหวัดประจวบคีรีขันธ์ มาลงบนระบบ e-Learning เพื่อใช้เผยแพร่แก่โรงเรียนในโครงการเทคโนโลยีสารสนเทศเพื่อการศึกษาของ โรงเรียนในชนบท หรือ ทสรช.ใช้ประโยชน์ในการสอน สอนเสริม หรือให้นักเรียนได้ใช้ทบทวนบทเรียนภายในโรงเรียนแบบ Off-line และเผยแพร่แบบ On-line ผ่านทางอินเทอร์เน็ต ให้แก่ครู นักเรียน และผู้สนใจทั่วไปได้ใช้ประโยชน์ในการเรียนการสอน หรือศึกษาเพิ่มเติม [5]

2.2 การพัฒนาหลักสูตร

ADDIE Model เป็นรูปแบบที่ได้รับการยอมรับในการนำมาใช้ในการออกแบบพัฒนาหลักสูตร แสดงดังแผนภาพที่ 1



ภาพที่ 1: ขั้นตอนรูปแบบ ADDIE Model

จากแผนภาพที่ 1 รูปแบบ ADDIE Model ประกอบด้วยขั้นตอนทั้งหมด 5 ขั้นตอนสามารถนำมาใช้ในการพัฒนาหลักสูตรดังนี้ 1) ขั้นตอนการวิเคราะห์ (Analysis) เป็นขั้นวางแผน วิเคราะห์ปัญหาหรือเตรียมการสื่อต่าง ๆ ที่จำเป็นต่อการพัฒนาหลักสูตร 2) ขั้นตอนการออกแบบ (Design) เป็นขั้นตอนการใช้ผลลัพธ์จากขั้นตอนการวิเคราะห์ เพื่อวางแผนสำหรับพัฒนาหลักสูตร ซึ่งจะต้องกำหนดโครงร่างวิธีการให้บรรลุถึงเป้าหมายการพัฒนาหลักสูตร 3) ขั้นตอนการพัฒนา (Development) เป็นขั้นตอนการพัฒนาสร้างชิ้นบนขั้นตอนการวิเคราะห์และการออกแบบจุดมุ่งหมายของขั้นตอนนี้คือ การสร้างหลักสูตรและสื่อทั้งหมดที่ใช้ในพัฒนาหลักสูตร 4) ขั้นตอนการทดลองใช้ (Implementation) เป็นขั้นที่นำหลักสูตรที่มีองค์ประกอบครบสมบูรณ์มาทดลองใช้เพื่อหาประสิทธิภาพของหลักสูตรการอบรม และ 5) ขั้นตอนประเมินผล (Evaluation) เป็นขั้นตอนสุดท้าย โดยการนำผลการทดลองที่ได้มาสรุปหาประสิทธิภาพของหลักสูตร ซึ่งกระบวนการทั้ง 5 ขั้นทำให้ได้หลักสูตรครบถ้วนเหมาะสมกับการอบรมได้เป็นอย่างดี [6]

2.3 งานวิจัยที่เกี่ยวข้อง

อรนัน ผันเฟื่อน [7] ทำวิจัยเรื่อง การพัฒนาหลักสูตรอบรมการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ เรื่อง การทำภาพสติกเกอร์ด้วย iPad2 สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 1 ผลการวิจัยสรุปได้ว่า 1) หลักสูตรอบรมที่พัฒนาขึ้นมีคุณภาพในระดับเหมาะสมมาก 2) ผู้เข้ารับการอบรมมีความรู้ความเข้าใจหลักสูตรอบรม โดยมีคะแนนหลังอบรมสูงกว่าก่อนอบรมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และ ผู้เรียนมีความพึงพอใจต่อหลักสูตรอบรมอยู่ในระดับมากที่สุด

นงเยาว์ ประธิปอาราม [8] ทำวิจัยเรื่อง การพัฒนาหลักสูตรอบรมการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ เรื่อง การตกแต่งรูปภาพโดยใช้ Ps Express ด้วย iPad 2 สำหรับชั้นมัธยมศึกษาปีที่ 4 ผลการวิจัยพบว่า 1) หลักสูตรอบรมการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ที่ผู้ศึกษาพัฒนาขึ้น มีคุณภาพในระดับมาก 2) ผู้เข้ารับการอบรมมีความรู้ความเข้าใจ

ในการอบรมโดยมีคะแนนหลังการอบรมสูงกว่าอบรมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 3) ผู้เข้ารับการอบรมมีความพึงพอใจต่อหลักสูตรการอบรมในระดับมาก

ประภาวดี กาญจนวิลานนท์ [9] ทำการวิจัยเรื่อง การพัฒนาหลักสูตรอบรมการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ เรื่อง การสร้างสื่อวีดิทัศน์ด้วย iPad 2 สำหรับชั้นมัธยมศึกษาปีที่ 6 ผลการวิจัยพบว่า 1) หลักสูตรอบรมการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ เรื่อง การสร้างสื่อวีดิทัศน์ iPad 2 มีคุณภาพในระดับมาก 2) ผู้เข้ารับการอบรมมีความรู้ ความเข้าใจในการอบรมโดยมีคะแนนหลังการอบรมสูงกว่าคะแนนก่อนการอบรมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 3) ผู้เข้ารับการอบรมมีความพึงพอใจต่อหลักสูตรการอบรมที่พัฒนาขึ้นในระดับมาก

นำชัย โบราณมูล [10] ทำวิจัยเรื่อง การพัฒนาหลักสูตรอบรมการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ เรื่อง การสร้างสเปรดชีตด้วย Numbers สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ผลการวิจัยพบว่า 1) หลักสูตรอบรมการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ที่ผู้ศึกษาพัฒนาขึ้น ประกอบไปด้วยเนื้อหา ความรู้พื้นฐานการใช้คอมพิวเตอร์แท็บเล็ต ส่วนประกอบของแท็บเล็ต วิธีการใช้แท็บเล็ต การใช้แอปพลิเคชัน การสร้างสเปรดชีตด้วย Numbers มีคุณภาพในระดับมาก 2) ผู้เข้ารับการอบรมมีความรู้ ความเข้าใจในการอบรมโดยมีคะแนนหลังการอบรมสูงกว่าก่อนการอบรมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และ 3) ผู้เข้ารับการอบรมมีความพึงพอใจต่อรูปแบบการอบรมในระดับมาก

สิริมณัส โยสีดา [11] ทำวิจัยเรื่อง การพัฒนาหลักสูตรอบรมการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ โดยใช้เทคนิคการอบรมกลุ่มร่วมมือแบบ STAD สำหรับครูสังกัดสำนักงานเขตพื้นที่การศึกษา เรื่อง การสร้างสเปรดชีตด้วยแท็บเล็ต ผลการวิจัยพบว่า หลักสูตรที่พัฒนาขึ้นประกอบด้วยจุดประสงค์การฝึกอบรม แนวทางการดำเนินการอบรมโดยใช้เทคนิคการอบรมกลุ่มร่วมมือ รูปแบบ STAD สื่อและคู่มือการวัดและประเมินผล ผลการประเมินคุณภาพหลักสูตรโดยผู้เชี่ยวชาญจำนวน 5 คน พบว่า หลักสูตรมีคุณภาพโดยรวมในระดับมากที่สุด ผลการทดลองใช้หลักสูตร พบว่าครูมีคะแนนเฉลี่ยหลังการอบรม แตกต่างจากคะแนนเฉลี่ยก่อนการอบรมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และครูมีความพึงพอใจต่อหลักสูตรในระดับมากที่สุด

3. ขั้นตอนวิธีการดำเนินการวิจัย

3.1 ขั้นตอนการวิจัย ผู้วิจัยดำเนินการตามรูปแบบ ADDIE ดังนี้

3.1.1 ขั้นการวิเคราะห์ ดำเนินการศึกษาเอกสารงานวิจัยที่เกี่ยวข้อง ศึกษาในด้านการพัฒนาหลักสูตร การประยุกต์ใช้แท็บเล็ตเพื่อการเรียนรู้ การพัฒนาแอปพลิเคชัน

3.1.2 ขั้นการออกแบบ ดำเนินการออกแบบหลักสูตรอบรมการประยุกต์ใช้สื่อ eDLTV เพื่อการเรียนรู้สู่การพัฒนาแอปพลิเคชันเกมเพื่อการเรียนรู้บนแท็บเล็ต และออกแบบเครื่องมือการประเมินผลต่าง ๆ

3.1.3 ขั้นการพัฒนา ดำเนินการพัฒนาหลักสูตรอบรมการประยุกต์ใช้สื่อ eDLTV เพื่อการเรียนรู้สู่การพัฒนาแอปพลิเคชันเกมเพื่อการเรียนรู้บนแท็บเล็ต และเครื่องมือการประเมินผลต่าง ๆ ตลอดจนหาคุณภาพของหลักสูตร และเครื่องมือต่าง ๆ

3.1.4 ขั้นทดลองใช้ ดำเนินการทดลองใช้กับกลุ่มเป้าหมายตามแบบแผนการวิจัย จัดเก็บข้อมูลโดยใช้เครื่องมือการวิจัย

3.1.5 ขั้นการประเมินผล ดำเนินการวิเคราะห์ และสรุปผลการวิจัย

3.2 กลุ่มเป้าหมายการวิจัย คือครูผู้สอนชั้นประถมศึกษาปีที่ 1 สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษามหาสารคาม เขต 1 จำนวน 30 คน ที่สมัครใจเข้ารับการอบรม

3.3 ตัวแปรการวิจัย ประกอบด้วย

ตัวแปรต้น ได้แก่ หลักสูตรอบรมครูการประยุกต์ใช้สื่อ eDLTV เพื่อการเรียนรู้สู่การพัฒนาแอปพลิเคชันเกมเพื่อการเรียนรู้บนแท็บเล็ต

ตัวแปรตาม ได้แก่ คุณภาพของหลักสูตรอบรมที่พัฒนาขึ้น ความรู้และความเข้าใจของครูที่เข้ารับการอบรม คุณภาพของแอปพลิเคชันเกมที่ครูพัฒนาขึ้น และความพึงพอใจของครูที่เข้ารับการอบรม

3.4 เครื่องมือการวิจัย ประกอบด้วย

3.4.1 หลักสูตร

3.4.2 แบบประเมินหลักสูตร

3.4.3 แบบทดสอบวัดผลการเรียนรู้

3.4.4 แบบประเมินคุณภาพแอปพลิเคชัน

3.4.5 แบบประเมินความพึงพอใจ

3.5 สถิติที่ใช้วิเคราะห์ข้อมูล ได้แก่ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และสถิติ t-test แบบ (Dependent Sample)

4. ผลการวิจัย

4.1 ผลการพัฒนาคุณภาพหลักสูตรอบรม

ผู้วิจัยได้พัฒนาหลักสูตรหลักสูตรอบรมการประยุกต์ใช้สื่ออีดีแอล เพื่อการเรียนรู้สู่การพัฒนาแอปพลิเคชันเกมเพื่อการเรียนรู้บนแท็บเล็ต ประกอบด้วยกระบวนการอบรมจำนวน 5 ขั้นตอน

ขั้นที่ 1 เรียนรู้จากการรับฟังการบรรยาย ในความสำคัญของสื่ออีดีแอลทีวีและแอปพลิเคชันเกมส์โดยใช้เครื่องมือของสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐานหรือ สพฐ.

ขั้นที่ 2 เรียนรู้จากการสาธิตการใช้สื่ออีดีแอลทีวี การเลือกเนื้อหา และสื่อที่ต้องการจากสื่ออีดีแอลทีวี และสาธิตการสร้างแอปพลิเคชันเกมส์โดยใช้เครื่องมือของสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐานหรือ สพฐ. จากเว็บไซต์ <http://www.otpcappcon.com>

ขั้นที่ 3 เรียนรู้การสร้างแอปพลิเคชันเกมส์ จากการศึกษาปฏิบัติการสร้างทีละขั้นตอน ดังนี้

- 1) เลือกเนื้อหา รูปภาพ จากสื่ออีดีแอลทีวี
- 2) ออกแบบเกมส์และเลือกองค์ประกอบของเกมส์จากเครื่องมือสร้างแอปพลิเคชันของ สพฐ.
- 3) ปฏิบัติการสร้างเกมส์ และทดสอบการใช้งาน

ขั้นที่ 4 เรียนรู้จากการให้คำแนะนำของวิทยากร และผู้ช่วยวิทยากรประเมินแอปพลิเคชันเกมส์ ที่พัฒนาขึ้น

ขั้นที่ 5 ปรับปรุงแก้ไขแอปพลิเคชันเกมส์ตามคำแนะนำของวิทยากร และผู้ช่วยวิทยากร

ผู้วิจัยได้นำหลักสูตรอบรมที่พัฒนาขึ้น นำเสนอผู้เชี่ยวชาญจำนวน 5 ท่าน ประกอบด้วยอาจารย์จากมหาวิทยาลัยราชภัฏมหาสารคาม ศึกษาพิเศษและครูจากสำนักงานเขตพื้นที่พิจารณาความเหมาะสมโดยใช้แบบประเมินคุณภาพหลักสูตรอบรมแบ่งออกเป็น 6 ด้าน ได้แก่ ด้านหลักการและเหตุผลด้านวัตถุประสงค์ของหลักสูตรอบรม ด้านโครงสร้างของหลักสูตรอบรม ด้านการจัดการ/กิจกรรมการอบรมด้านการวัดผลการอบรมด้านการประเมินผลการอบรม ผลการประเมินแสดงในตารางที่ 1

ตารางที่ 1: ผลการประเมินคุณภาพ

รายการ	\bar{X}	S.D.	ความหมาย
1. หลักการและเหตุผล	4.57	0.52	มากที่สุด
2.วัตถุประสงค์ของหลักสูตรอบรม	4.67	0.49	มากที่สุด
3. โครงสร้างของหลักสูตรอบรม	4.55	0.51	มากที่สุด
4.ด้านการจัดการ/กิจกรรมการอบรม	4.60	0.50	มากที่สุด
5.การวัดผลการอบรม	4.50	0.53	มากที่สุด
6.การประเมินผลการอบรม	4.50	0.53	มากที่สุด
ความเหมาะสมเฉลี่ยโดยรวม	4.57	0.49	มากที่สุด

จากตารางที่ 1 พบว่า หลักสูตรอบรมการประยุกต์ใช้สื่อ eDLTV เพื่อการเรียนรู้สู่การพัฒนาแอปพลิเคชันเกมเพื่อการเรียนรู้บนแท็บเล็ต ผู้เชี่ยวชาญมีความคิดเห็นโดยรวมในระดับเหมาะสมมากที่สุด ($\bar{X} = 4.57$ และ S.D.= 0.49)

4.2 ผลการศึกษาความรู้และความเข้าใจ

ผู้วิจัยทดสอบความรู้และความเข้าใจของครูจำนวน 30 คน ทั้งก่อนอบรมและหลังอบรมด้วยแบบทดสอบวัดความรู้และความเข้าใจ และนำคะแนนที่ได้ มาวิเคราะห์คำนวณค่าสถิติ t-test ผลการวิเคราะห์แสดงในตารางที่ 2

ตารางที่ 2 : ผลการวิเคราะห์

ครู	คะแนนเต็ม	ก่อนอบรม	หลังอบรม	t
30	30	13.23	29.29	29.71

จากตารางที่ 2 พบว่าคะแนนเฉลี่ยก่อนอบรมมีค่า 13.23 คะแนนเฉลี่ยหลังอบรมมีค่า 29.29 เมื่อเปรียบเทียบค่า t ที่คำนวณได้มีค่า 29.71 ซึ่งมีค่ามากกว่าค่า t (ตาราง $\alpha = .05, df = 29$) (1.699) สรุปได้ว่า คะแนนความรู้และความเข้าใจเฉลี่ยหลังการอบรมสูงกว่าก่อนการอบรม อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

4.3 ผลการศึกษาคุณภาพพัฒนาแอปพลิเคชันของผู้เข้ารับการอบรม

ผู้วิจัยนำแอปพลิเคชัน จากครูผู้เข้ารับการอบรมตามหลักสูตรที่พัฒนาขึ้น จำนวน 30 ชิ้นงาน ประเมินคุณภาพโดยวิทยากรและผู้ช่วยวิทยากร รวม 5 คน ผลการประเมินคุณภาพ พบว่าแอปพลิเคชันที่พัฒนาขึ้นมีคุณภาพอยู่ในระดับมาก ถึง มากที่สุด (\bar{X} อยู่ระหว่าง 4.25 – 4.83, S.D.อยู่ระหว่าง 0.39-0.90)

4.4 ผลการศึกษาความพึงพอใจของผู้เข้ารับการอบรม

ผู้วิจัยได้ศึกษาความพึงพอใจของผู้เรียนหลังจากได้จัดการอบรมหลักสูตรที่พัฒนาขึ้น ผลการประเมินแสดงในตารางที่ 3

ตารางที่ 3: ผลการศึกษาความพึงพอใจ

รายการ	\bar{X}	S.D.	ความหมาย
1. ด้านวิทยากร	4.55	0.11	มากที่สุด
2. ด้านกระบวนการอบรม	4.65	0.29	มากที่สุด
3. ด้านความรู้ความเข้าใจ	4.85	0.21	มากที่สุด
4. ด้านการนำไปใช้ประโยชน์	4.83	0.30	มากที่สุด
5. ด้านสถานที่ระยะเวลา	4.36	0.39	มาก
ความพึงพอใจเฉลี่ยโดยรวม	4.65	0.41	มากที่สุด

จากตารางที่ 3 พบว่า ผู้เข้ารับการอบรมมีความพึงพอใจต่อการจัดการอบรมด้วยหลักสูตรที่พัฒนาขึ้น โดยภาพรวมเฉลี่ยทั้ง 5 ด้าน พบว่า ผู้เข้ารับการอบรมส่วนใหญ่มีความพึงพอใจอยู่ในระดับมากที่สุด ($\bar{X}=4.65, S.D.=0.41$)

5. สรุปผลการวิจัย

1. ผลการพัฒนาหลักสูตรอบรมประกอบด้วยกระบวนการอบรม 5 ขั้นตอน ขั้นที่ 1 เรียนรู้จากการรับฟังการบรรยาย ขั้นที่ 2 เรียนรู้จากการสาธิตการใช้สื่ออีดีแอลทีวี และสาธิตการสร้างแอปพลิเคชันเกมส์ ขั้นที่ 3 เรียนรู้การสร้างแอปพลิเคชันเกมส์ จาก การปฏิบัติการ ขั้นที่ 4 เรียนรู้จากการให้คำแนะนำของวิทยากร และผู้ช่วยวิทยากร ขั้นที่ 5 ปรับปรุงแก้ไขแอปพลิเคชันเกมส์ และผู้เชี่ยวชาญ 5 ท่าน มีความคิดเห็น โดยรวมในระดับเหมาะสมมาก

2. ผลการศึกษาความรู้ความเข้าใจ สรุปได้ว่าคะแนน ความรู้และความเข้าใจหลังการอบรมสูงกว่าก่อนการอบรม อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

3. ผลการศึกษาคูณภาพแอปพลิเคชันของครูผู้เข้ารับการ อบรม พบว่าแอปพลิเคชันที่พัฒนาขึ้น มีคุณภาพอยู่ในระดับ มาก ถึง มากที่สุด

4. ผลการศึกษาคความพึงพอใจของผู้เข้ารับการอบรมพบว่า โดยรวมมีความพึงพอใจอยู่ในระดับมากที่สุด

6. เอกสารอ้างอิง

[1] สำนักงานเลขาธิการรัฐมนตรี .“โครงการแท็บเล็ตเพื่อการ ศึกษา” สืบค้นเมื่อ จาก <http://suksa.2moe.go.th> (เมื่อ 10 มกราคม 2556).

[2] สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน “คู่มือการ อบรมการสร้างแอปพลิเคชัน 2556 ”
สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน.

[3] สมเจตน์ ภูศรี และ วรปภา อารีราษฎร์ “รายงานการ

ดำเนินงานเครือข่ายเผยแพร่ถ่ายทอดและพัฒนาสื่อการ เรียนการสอนบนระบบ e-Learning (eDL-Square)”.
มหาวิทยาลัยราชภัฏ มหาสารคาม, 2555

[4] ราชภัฏมหาสารคาม “มหาวิทยาลัยราชภัฏมหาสารคาม ดำเนินงานการส่งเสริมการพัฒนาครู บุคลากรทางการศึกษา และนักเรียนในการใช้เทคโนโลยีสารสนเทศเพื่อ ยกระดับการศึกษาเรียนรู้ของชุมชนมหาวิทยาลัยราชภัฏมหาสารคาม” 2555.

[5] โครงการ “จัดทำเนื้อหา ระบบ e-Learning ของการศึกษาทางไกล ผ่านดาวเทียม เฉลิมพระเกียรติเนื่องในโอกาสสมทวงกล เฉลิมพระ ชนมพรรษา 80 พรรษา 5 ธันวาคม 2550. (โครงการ eDLTV)” [ออนไลน์].สืบค้นจาก <http://edltv.thai.net> (เมื่อ 12 มีนาคม 2556) .

[6] พิสุทธิ อารีราษฎร์. “การพัฒนาซอฟต์แวร์ทางการศึกษา. มหาสารคาม” มหาวิทยาลัยราชภัฏมหาสารคาม, 2551.

[7] อรฉัตร เตือนพั้ง “การพัฒนาหลักสูตรอบรมการใช้ คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ เรื่อง การทำภาพสติกเกอร์ ด้วย iPad2” มหาวิทยาลัยราชภัฏมหาสารคาม 2555.

[8] นงเยาว์ ประธิปอราม “การพัฒนาหลักสูตรอบรมการใช้ คอมพิวเตอร์แท็บเล็ต เพื่อการเรียนรู้ เรื่องการตกแต่งรูปภาพโดย ใช้ Ps Express ด้วย iPad2” มหาวิทยาลัยราชภัฏมหาสารคาม, 2555.

[9] ประภาวดี กาญจนวิลานนท์. “การพัฒนาหลักสูตรอบรมการใช้ คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ เรื่อง การสร้างสื่อวีดิทัศน์ด้วย iPad2 สำหรับชั้นมัธยมศึกษาปีที่ 6” มหาวิทยาลัยราชภัฏ มหาสารคาม 2555.

[10] นาชัย โบราณมูล. “การพัฒนาหลักสูตรอบรมการใช้คอมพิวเตอร์ แท็บเล็ตเพื่อการเรียนรู้ เรื่องการสร้างสปรดชีทด้วย Numbers สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3” มหาวิทยาลัยราชภัฏ มหาสารคาม 2555.

[11] สิริมนัส โยธิตา “การพัฒนาหลักสูตรอบรมการใช้คอมพิวเตอร์แท็บ เล็ตเพื่อการเรียนรู้ โดยใช้เทคนิคการอบรมกลุ่มร่วมมือแบบ STAD เรื่อง การสร้างสปรดชีทด้วยแท็บเล็ต “ มหาวิทยาลัยราช ภัฏมหาสารคาม 2555.

ผลการสังเคราะห์รูปแบบกิจกรรมค่ายอาสาเพื่อการเรียนรู้สื่ออีดีแอลทีวี

สำหรับโรงเรียนขนาดเล็ก เครือข่ายมหาวิทยาลัยราชภัฏมหาสารคาม

Guidelines on Volunteer Camp Activity Model for Learning eLDTV Media of Small School of Rajabhat Maha Satrakham University Network

อภิชาติ เหล็กดี (Apichat Lagdee)¹ สมเจตน์ ภูศรี (Somjet Poosri)² วรภา อารีราษฎร์ (Worapapha Arreerard)³

¹นักศึกษาลัทธิศึกษาศาสตร์บัณฑิต มหาวิทยาลัยราชภัฏมหาสารคาม ²มหาวิทยาลัยราชภัฏมหาสารคาม

³คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏมหาสารคาม

¹pugan31@gmail.com, ²somjet.p@rnu.mail.go.th, ³dr.worapapha@hotmail.com

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ 1) ศึกษาความต้องการของครูในการจัดกิจกรรมค่ายอาสาเพื่อการเรียนรู้สื่ออีดีแอลทีวี 2) สังเคราะห์รูปแบบกิจกรรมค่ายอาสาเพื่อการเรียนรู้สื่ออีดีแอลทีวี เครือข่ายมหาวิทยาลัยราชภัฏมหาสารคาม และ 3) สอบถามความคิดเห็นของผู้เชี่ยวชาญที่มีต่อรูปแบบที่สังเคราะห์ขึ้น กลุ่มตัวอย่างที่ใช้ในการวิจัยเป็น ผู้บริหารและผู้ประสานงาน โรงเรียนเครือข่ายสื่ออีดีแอลทีวี 111 โรงเรียน เครื่องมือที่ใช้ในการวิจัย คือ แบบสำรวจความต้องการจัดกิจกรรมค่ายอาสาและแบบสอบถามความคิดเห็นของผู้เชี่ยวชาญที่มีต่อรูปแบบ และสถิติที่ใช้คือ ร้อยละ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน

ผลการวิจัยพบว่า 1) ผลการศึกษาความต้องการของครูในการจัดกิจกรรมค่ายอาสาเพื่อการเรียนรู้สื่ออีดีแอลทีวี พบว่า 1.1) ด้านรูปแบบการเข้าร่วมกิจกรรมประกอบด้วย 4 กิจกรรม คือ ก) กิจกรรมแบบกลุ่มโรงเรียน ร้อยละ 52.25 ข) กิจกรรมแบบแยกโรงเรียน ร้อยละ 33.33 ค) กิจกรรมแบบเวิร์กช็อป ร้อยละ 3.60 และ ง) กิจกรรมแบบนักศึกษาฝึกสอน ร้อยละ 1.80 และ 1.2) ด้านความต้องการของโรงเรียนในการจัดกิจกรรมค่ายอาสาประกอบด้วย ก) ความต้องการเรียนรู้วิธีการใช้งานสื่ออีดีแอลทีวี พบว่า มีความต้องการแบบออนไลน์-ออฟไลน์ ร้อยละ 98.20 การใช้งานแบบระบบเครือข่ายภายในโรงเรียน ร้อยละ 84.68 การใช้งานแบบเวิร์กช็อป ร้อยละ 3.60 และการใช้งานผ่านคอมพิวเตอร์แท็บเล็ต ร้อยละ 86.49 และ ข) ความต้องการกิจกรรมเสริมการออกค่ายอาสา พบว่า ต้องการเรียนรู้การดำเนินสื่อ ร้อยละ 88.29 การซ่อมบำรุง/เครือข่าย ร้อยละ 57.66 และการพัฒนาหนังสืออิเล็กทรอนิกส์ ร้อยละ 30.63 2) ผลการสังเคราะห์รูปแบบการจัดกิจกรรมค่ายอาสาเพื่อการเรียนรู้สื่ออีดีแอลทีวีสำหรับโรงเรียนขนาดเล็กพบว่า กิจกรรมมีองค์ประกอบ 7 ส่วน คือ 2.1) นโยบายภาครัฐ 2.2) สื่ออีดีแอลทีวี 2.3)

โรงเรียนขนาดเล็ก 2.4) การเรียนรู้สื่ออีดีแอลทีวี 2.5) กิจกรรมค่ายอาสา 2.6) หน่วยงานความร่วมมือ และ 2.7) ตัวชี้วัด 5 ตัว และ 3) ความคิดเห็นของผู้เชี่ยวชาญที่มีต่อผลการสังเคราะห์รูปแบบการจัดกิจกรรมค่ายอาสาเพื่อการเรียนรู้สื่ออีดีแอลทีวีอยู่ในระดับเหมาะสมมากที่สุด ($\bar{X} = 4.65, S.D. = 0.05$)

คำสำคัญ: กิจกรรมค่ายอาสา การเรียนรู้สื่ออีดีแอลทีวี โรงเรียนขนาดเล็ก เครือข่ายมหาวิทยาลัยราชภัฏมหาสารคาม

Abstract

The purposes of this research were 1) survey needs of teachers for organizing volunteer camp activities for learning eDLTV media, 2) to synthesize models of volunteer camp activities model for learning eDLTV media of small school networks of Rajabhat MahaSarakhm University, and 3) to survey opinions of the experts on an improved model for learning eDLTV media. The samples used in this research were administrators and coordinators from 111 eDLTV Media school networks. The instrument was two sets of questionnaires for school administrators including coordinators and experts.. The statistics used were percentage, mean and standard deviation.

The research results showed that: 1) The eLDTV learning activities obviously needed by the teachers consisted of four major activities: school group activities(52.25%),single school activities (33.33%), server activities (3.60%), and internship training activities (1.80%). 2) The volunteer camp activities really needed by school networks included online – offline learning program (98.20%), school networks (LAN)

(84.68%), using server (3.60%) and Tablet applications (86.49%). 3) The volunteer camp activities obviously needed by the teachers consisted of learning media copying (88.29%), maintenance / networks (57.66%) and e-book development (30.63%). 4) Regarding the synthesis of volunteer camp activities model, the study showed that the eLDTV learning activities for small schools included 7 factors: government policies, eLDTV media, school size, learning eLDTV media, volunteer camp activities, collaborative organizations and five indicators. 4) The finding showed that the average level of the opinions of the experts on the appropriateness of a volunteer camp activities model for learning eDLTV media was very high ($X = 4.65$, $S.D. = .05$).

Keywords: Volunteer Camp Activities, Learning eDLTV Media , Small Schools , Rajabhat Maha Sarakham University Networks

1. บทนำ

สื่ออีดีแอลทีวี [3] หรือ eDLTV (Electronic Distance Learning via Television) เป็นสื่อการเรียนการสอนที่เป็นผลจากการดำเนินงานของโครงการจัดทำเนื้อหาในระบบ e-Learning ของการศึกษาทางไกลผ่านดาวเทียมเฉลิมพระเกียรติเนื่องในโอกาสมหามงคลเฉลิมพระชนมพรรษา 5 ธันวาคม 2550 ภายใต้ความร่วมมือของมูลนิธิการศึกษาทางไกลผ่านดาวเทียมร่วมกับโครงการเทคโนโลยีสารสนเทศตามพระราชดำริสมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี



ภาพที่ 1 : จอภาพระบบอีดีแอลทีวี

โดยได้คัดสรรบทเรียนที่มีคุณภาพที่ออกอากาศทางสถานีโทรทัศน์การศึกษาทางไกลผ่านดาวเทียม จากโรงเรียนวังไกลกังวล ประกอบด้วย สื่อการสอนระดับประถมศึกษา มัธยมศึกษา และอาชีวศึกษากลุ่มการทำงานอาชีพ เพื่อเผยแพร่ทั้งแบบออฟไลน์และแบบออนไลน์ให้แก่ครู นักเรียน และผู้สนใจทั่วไปได้ใช้ประโยชน์ในการเรียนการสอน หรือศึกษาเพิ่มเติม [3]

มหาวิทยาลัยราชภัฏมหาสารคามเป็นสถาบันการศึกษาเพื่อพัฒนาท้องถิ่น ได้เข้าร่วมเป็นสมาชิกเครือข่ายเผยแพร่ ถ่ายทอด และพัฒนาสื่อการเรียนการสอนบนระบบอีดีแอลทีวี หรือการเผยแพร่สื่ออีดีแอลทีวี ในการเรียนการสอนให้กับโรงเรียนในเขตพื้นที่บริการในปี พ.ศ. 2552 เป็นต้นมา [5] การดำเนินงาน พ.ศ. 2555-2556 มหาวิทยาลัยได้นำนโยบายสู่การปฏิบัติร่วมกับสมาชิกเครือข่าย ภายใต้การดำเนินงาน 5 กิจกรรม คือ 1) การเผยแพร่การใช้งานอีดีแอลทีวี อีดีแอลอาร์ยู อาร์เอ็มยูอีดีแอล 2) การเผยแพร่นวัตกรรมอีดีแอลทีวี ให้เป็นประโยชน์ต่อการจัดการศึกษาของนักศึกษา 3) การให้บริการวิชาการแก่โรงเรียนและชุมชนในท้องถิ่นอบรมประยุกต์ใช้สื่ออีดีแอลทีวีในการเรียนการสอน และ 4) การจัดกิจกรรมค่ายอาสาเยาวชนทั่วถิ่นไทย เรียนรู้ได้ร่วมพระบารมี (เกิดพระเกียรติ 60 พระชันษามหาจักรี สิริินทร) และ 5) การแลกเปลี่ยนเรียนรู้และสรุปผลการดำเนินงาน [5]

กิจกรรมค่ายอาสา เป็นกิจกรรมที่มหาวิทยาลัยราชภัฏมหาสารคาม จัดร่วมกับโรงเรียนในเครือข่าย โดยจัดกิจกรรมที่หลากหลาย เพื่อมุ่งเน้นการส่งเสริมการเรียนรู้สื่ออีดีแอลทีวี ให้ครูโรงเรียนได้เรียนรู้และเข้าใจ สามารถประยุกต์ใช้ในกิจกรรมการเรียนการสอนได้อย่างมีประสิทธิภาพ การจัดกิจกรรมขึ้นอยู่กับความพร้อมและความต้องการของโรงเรียน [8] ผลการจัดกิจกรรมค่ายอาสาในระยะ พบว่า โรงเรียนต้องการให้เพิ่มเวลาในการจัดกิจกรรม ให้กับโรงเรียนที่มีบริบทของโรงเรียนที่แตกต่างกัน เช่น ครูไม่ครบชั้น ครูไม่ตรงสาย ครูสอนไม่ตรงกลุ่มสาระ สื่ออีดีแอลทีวีสามารถช่วยแก้ปัญหาการเรียนการสอนได้ โรงเรียนยังต้องการคำแนะนำด้านการบริหารจัดการวิธีการใช้งาน และการประยุกต์ใช้สื่ออีดีแอลทีวีที่สอดคล้องกับบริบทของโรงเรียน รวมถึงการส่งสื่อให้นักเรียนไปเรียนรู้ที่บ้านตามนโยบายห้องเรียนกลับด้าน [6]

จากการศึกษาเอกสาร การดำเนินการเผยแพร่ และถ่ายทอดสื่ออีดีแอลทีวี ของมหาวิทยาลัยและการศึกษาบริบทของโรงเรียนขนาดเล็ก ผู้วิจัยมีความสนใจศึกษา กิจกรรมค่ายอาสาเพื่อการเรียนรู้สื่ออีดีแอลทีวีของโรงเรียนขนาดเล็ก ซึ่งการดำเนินงานเผยแพร่สื่ออีดีแอลทีวีสู่โรงเรียนขนาดเล็ก ให้เกิดประโยชน์ต่อการเรียนรู้ของผู้เรียน การเชื่อมความสัมพันธ์ระหว่างโรงเรียนกับชุมชนด้วยสื่ออีดีแอลทีวี จึงเป็นกิจกรรมการสร้างเครือข่ายการเรียนรู้ของมหาวิทยาลัยราชภัฏมหาสารคาม ร่วมกับ โรงเรียน และชุมชน จะนำมาสู่การส่งเสริมให้โรงเรียนขนาดเล็กมีสื่ออีดีแอลทีวี กิจกรรมค่ายอาสาจะเป็นกิจกรรมหนึ่งที่ส่งเสริมให้ ผู้บริหาร ครู และนักเรียน เรียนรู้การประยุกต์ใช้สื่ออีดีแอลทีวีเพื่อการเรียนการสอน

สอน ผู้วิจัยจึงได้ดำเนินการสังเคราะห์รูปแบบการจัดกิจกรรม ค่ายอาสาเพื่อการเรียนรู้สื่ออีดีแอลทีวี ที่สอดคล้องและเหมาะสมกับบริบทของโรงเรียนขนาดเล็ก เพื่อนำไปใช้เป็นเครื่องมือในการวิจัยในระยะต่อไป

2. วรรณกรรมที่เกี่ยวข้อง

2.1 ความสำคัญของสื่ออีดีแอลทีวี

ระบบอีดีแอลทีวี [3] หรือ สื่ออีดีแอลทีวี เป็นสื่ออิเล็กทรอนิกส์ที่ประกอบด้วย วิกิพีเดียการสอน สไลด์ ใบงาน ใบความรู้ และแบบทดสอบ รวบรวมและจัดเก็บด้วยระบบบริหารจัดการเรียนรู้อีดีแอลสแควร์ (eDL-Square) สามารถเข้าถึงเนื้อหาได้ทั้งระบบออนไลน์และระบบออฟไลน์ สื่ออีดีแอลทีวีมี 4 ระบบคือ สื่อระดับปฐมวัย ประถมศึกษา มัธยมศึกษา และสื่อเพื่อการทำงานอาชีพ โดยมีสื่ออิเล็กทรอนิกส์ รวมทั้งสิ้น 73,867 ชิ้น มหาวิทยาลัยราชภัฏมหาสารคาม ได้ทำการเผยแพร่สื่ออีดีแอลทีวีทั้ง 4 ระบบ โดยบรรจุในฮาร์ดดิสก์ ขนาด 2 เทราไบต์ พร้อมกับจัดกิจกรรมค่ายอาสาเพื่อส่งเสริมการเรียนรู้สื่ออีดีแอลทีวีอย่างต่อเนื่อง [6] จากผลการดำเนินงาน การเผยแพร่ ถ่ายทอด และพัฒนาสื่ออีดีแอลทีวี ด้วยรูปแบบการพัฒนาสื่ออิเล็กทรอนิกส์สู่การวิจัยในชั้นเรียน [8] พบว่ารูปแบบที่พัฒนาขึ้น ประกอบด้วย กระบวนการ 5 ขั้นตอน คือ 1) การสร้างความตระหนัก 2) การส่งเสริมการเรียนรู้ 3) การสร้างสื่ออิเล็กทรอนิกส์ 4) การวิจัยในชั้นเรียน และ 5) การติดตามประเมินผล ผลการนำรูปแบบไปใช้พบว่าผู้เข้ารับการอบรมตามรูปแบบมีความพึงพอใจระดับมากที่สุด

2.2 การเรียนรู้แบบกิจกรรมค่าย

กิจกรรมค่าย [1] เป็นกิจกรรมการเรียนรู้ ที่จัดอยู่ในลักษณะของการเปลี่ยนผ่านการศึกษา (Education Transformation) เป็นการเปลี่ยนแนวทางการจัดการศึกษา เพื่อพัฒนาผู้เรียนในสังคมวัฒนธรรมร่วมสมัย ให้เป็นผู้ที่มีศักยภาพในการเรียนรู้ทั้งเป็นรายบุคคลและเรียนรู้ร่วมกัน โดยใช้แนวคิดความร่วมมือขององค์กรต่างๆ ในสังคม เพื่อให้เกิดผลอันพึงประสงค์ กิจกรรมค่ายจึงเป็นกิจกรรมการเรียนรู้ ที่จัดอยู่ในประเด็น การสร้างโอกาสการเรียนรู้ ที่ครูและผู้บริหารควรสร้างความตระหนักให้นักเรียนเห็นความสำคัญ มหาวิทยาลัยราชภัฏมหาสารคาม [7] ได้จัดกิจกรรมค่ายร่วมกับโรงเรียนเครือข่าย ประกอบด้วย 3 กิจกรรม คือ 1) กิจกรรมการเรียนรู้สื่ออีดีแอลทีวีกับการพัฒนาหนังสืออิเล็กทรอนิกส์ 2) กิจกรรมการเรียนรู้สื่ออีดีแอลทีวีกับคอมพิวเตอร์แท็บเล็ต และ 3) กิจกรรมการเรียนรู้สื่ออีดีแอลทีวีกับฮาร์ดแวร์ โดยจัดกิจกรรมที่โรงเรียน 1 วัน ผู้เข้าร่วมกิจกรรมมีความพึงพอใจต่อการจัดกิจกรรมในระดับมากที่สุด

3. วิธีการดำเนินการวิจัย

3.1 ประชากรและกลุ่มตัวอย่าง

3.1.1 ประชากรที่ใช้ในการวิจัยครั้งนี้ ประกอบด้วย

1) ประชากรและกลุ่มตัวอย่างสำหรับการสำรวจข้อมูล เป็นบุคลากรโรงเรียนขนาดเล็กที่ได้รับสื่ออีดีแอลทีวีจากโครงการ จำนวน 111 แห่ง คัดเลือกเป็นกลุ่มตัวอย่างโดยวิธีเจาะจง โรงเรียนละ 2 คน รวมทั้งสิ้น 222 คน ประกอบด้วยผู้บริหารโรงเรียนละ 1 คน และ ผู้ประสานงานโรงเรียนละ 1 คน

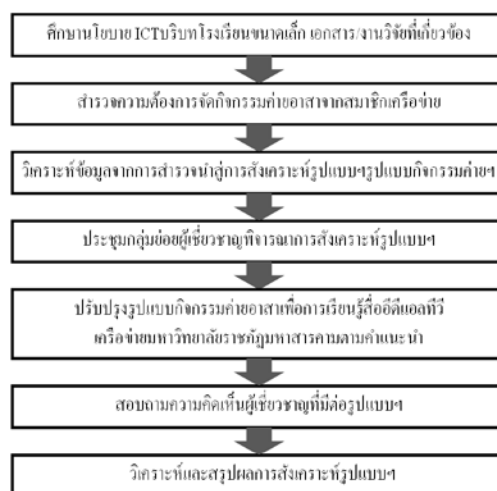
2) กลุ่มเป้าหมาย ในการประเมินรูปแบบที่สังเคราะห์ขึ้นเป็นผู้เชี่ยวชาญที่มีประสบการณ์ใช้สื่ออีดีแอลทีวี จำนวน 9 คน จาก 3 กลุ่ม กลุ่มละ 3 คน คือ กลุ่มนักวิชาการ กลุ่มผู้บริหาร และกลุ่มศึกษานิเทศก์และครูผู้สอน

3.2 เครื่องมือที่ใช้ในการวิจัย

3.2.1 แบบสำรวจความต้องการจัดกิจกรรมค่ายอาสาเพื่อการเรียนรู้สื่ออีดีแอลทีวี

3.2.2 แบบสอบถามความคิดเห็นของผู้เชี่ยวชาญที่มีต่อการสังเคราะห์กิจกรรมค่ายอาสาเพื่อการเรียนรู้สื่ออีดีแอลทีวี

3.3 ขั้นตอนการดำเนินงาน



ภาพที่ 2 : ขั้นตอนการดำเนินงานวิจัย

จากภาพที่ 2 ขั้นตอนการดำเนินงานวิจัย เริ่มจากการศึกษานโยบาย การสำรวจความต้องการ วิเคราะห์ข้อมูลจากผลการสำรวจ นำสู่การสังเคราะห์รูปแบบการจัดกิจกรรมค่ายอาสาฯ จัดประชุมกลุ่มย่อยผู้เชี่ยวชาญพิจารณาการสังเคราะห์รูปแบบ ปรับปรุงรูปแบบ แล้วนำเสนอต่อผู้เชี่ยวชาญเพื่อสอบถามความคิดเห็นที่มีต่อผลการสังเคราะห์รูปแบบฯ นำผลการสอบถามผู้เชี่ยวชาญมาวิเคราะห์และสรุปผลการวิจัย

3.4 การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูล ผู้วิจัยนำแบบสำรวจที่ได้จากกลุ่มตัวอย่าง มาวิเคราะห์ระดับความต้องการโดยใช้สถิติ ร้อยละ และ

วิเคราะห์ความคิดเห็นของผู้เชี่ยวชาญ ด้วยสถิติ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน นำค่าเฉลี่ยที่ได้มาเทียบกับเกณฑ์ ดังนี้ [4]

- 4.51 – 5.00 หมายถึง เหมาะสมมากที่สุด
- 3.51 – 4.50 หมายถึง เหมาะสมมาก
- 2.51 – 3.50 หมายถึง เหมาะสมปานกลาง
- 1.51 – 2.50 หมายถึง เหมาะสมน้อย
- 1.00 – 1.50 หมายถึง เหมาะสมน้อยที่สุด

4. ผลการดำเนินงาน

4.1 ผลการสำรวจความต้องการจัดกิจกรรมค่ายอาสาเพื่อการเรียนรู้สื่ออีดีแอลทีวี แสดงดังตารางที่ 1

ตารางที่ 1 ผลการสำรวจความต้องการจัดกิจกรรมค่ายอาสา

รายการ	ร้อยละ
1. ด้านรูปแบบการเข้าร่วมกิจกรรม	
1.1 กิจกรรมแบบกลุ่มโรงเรียน	52.25
1.2 กิจกรรมแบบแยกโรงเรียน	33.33
1.3 กิจกรรมแบบเวิร์กช็อป	3.60
1.4 กิจกรรมแบบนักศึกษาศึกษาประสบการณ์วิชาชีพครู	1.80
2. ความต้องการของโรงเรียนในการจัดกิจกรรมค่าย	
2.1 ด้านความต้องการเรียนรู้สื่ออีดีแอลทีวี	98.20
ก. การใช้งานแบบออนไลน์-ออฟไลน์	84.68
ข. การใช้งานผ่านระบบเครือข่ายภายในของโรงเรียน	
ค. การใช้งานแบบเวิร์กช็อป	3.60
ง. การใช้งานผ่านแท็บเล็ต	86.49
2.2 ความต้องการกิจกรรมเสริมการออกค่าย	
ก. การเรียนรู้การสำเนาสื่ออีดีแอลทีวี	88.29
ข. การซ่อมบำรุงระบบและเครือข่าย	57.66
ค. การพัฒนาหนังสืออิเล็กทรอนิกส์	30.63

จากตารางที่ 1 ผลการสำรวจข้อมูลพบว่า

1. ด้านรูปแบบการเข้าร่วมกิจกรรมประกอบด้วย 4 กิจกรรมคือ 1.1) กิจกรรมแบบกลุ่มโรงเรียน ร้อยละ 52.25 1.2) กิจกรรมแบบแยกโรงเรียน ร้อยละ 33.33 1.3) กิจกรรมแบบเวิร์กช็อป ร้อยละ 3.60 และ 1.4) กิจกรรมแบบนักศึกษาศึกษาประสบการณ์วิชาชีพครู ร้อยละ 1.80

2. ด้านความต้องการของโรงเรียนในการจัดกิจกรรมค่าย จำแนกเป็น 2 ด้าน 2.1) ความต้องการเรียนรู้สื่ออีดีแอลทีวี พบว่า ก) การใช้งานแบบออนไลน์-ออฟไลน์ ร้อยละ 84.20 ข) การใช้งานผ่านระบบเครือข่ายภายในโรงเรียน ร้อยละ 84.68 ค) การใช้งานแบบเวิร์กช็อป ร้อยละ 3.60 และ ง) การใช้งานผ่านแท็บเล็ต ร้อยละ 86.49 และ 2.2) ความต้องการกิจกรรมเสริมขณะออกค่ายอาสา พบว่า ก) การเรียนรู้การสำเนาสื่ออีดีแอล

ทีวี ร้อยละ 88.29 ข) การซ่อมบำรุงระบบและเครือข่าย ร้อยละ 57.66 และ ค) การพัฒนาหนังสืออิเล็กทรอนิกส์ ร้อยละ 30.63

4.2 ผลการสังเคราะห์รูปแบบการจัดกิจกรรมค่ายอาสาเพื่อการเรียนรู้สื่ออีดีแอลทีวี สำหรับโรงเรียนขนาดเล็ก มหาวิทยาลัยราชภัฏมหาสารคาม

1) นิยามรูปแบบ คือ องค์ประกอบของการจัดกิจกรรมค่ายอาสาเพื่อการเรียนรู้สื่ออีดีแอลทีวี ที่นำนโยบายภาครัฐด้านเทคโนโลยีสารสนเทศและการสื่อสาร และความสำคัญของสื่อไอซีที มาประยุกต์ใช้สื่ออีดีแอลทีวีเพื่อการเรียนการสอน ให้สอดคล้องกับบริบทของโรงเรียนและสอดคล้องกับความต้องการของครู โดยอาศัยกิจกรรมค่ายอาสาและความร่วมมือของหน่วยงานทั้งภาครัฐและเอกชน

2) องค์ประกอบของรูปแบบ มี 7 ส่วน ดังภาพที่ 3



ภาพที่ 3 : รูปแบบการจัดกิจกรรมค่ายอาสาเพื่อการเรียนรู้สื่ออีดีแอลทีวี สำหรับโรงเรียนขนาดเล็ก

จากภาพที่ 3 องค์ประกอบแต่ละส่วนมีรายละเอียด ได้ดังนี้

2.1 นโยบายภาครัฐ เป็นหน่วยงานกำหนดนโยบายด้านเทคโนโลยีสารสนเทศและการสื่อสาร หรือ ไอซีทีเพื่อการเรียนรู้ ประกอบด้วย กรอบนโยบายเทคโนโลยีสารสนเทศและการสื่อสาร ระยะ พ.ศ. 2554-2563 ของประเทศไทย (IT 2020) แผนแม่บทเทคโนโลยีสารสนเทศและการสื่อสาร (ฉบับที่ 2) ของประเทศไทย พ.ศ. 2552-2556 แผนแม่บทเทคโนโลยีสารสนเทศและการสื่อสารเพื่อการศึกษากระทรวงศึกษาธิการ พ.ศ. 2554 – 2556 ข้อเสนอการปฏิรูปการศึกษาในทศวรรษที่สอง (พ.ศ.2552-2556) นโยบายและภารกิจของมหาวิทยาลัยราชภัฏมหาสารคาม การให้บริการวิชาการแก่ชุมชนและท้องถิ่น

2.2 สื่ออีดีแอลทีวี เป็นสื่อไอซีที การส่งเสริมครูประยุกต์ใช้สื่ออีดีแอลทีวีด้วยกิจกรรมค่ายอาสา ภายใต้บริบทของโรงเรียนและความต้องการของครูเป็นการนำสู่การปฏิบัติที่โรงเรียน

2.3 โรงเรียนขนาดเล็ก เป็นโรงเรียนสมาชิกเครือข่ายที่ได้รับสื่ออีดีแอลทีวีเพื่อจัดกิจกรรมค่ายอาสา ร่วมกับมหาวิทยาลัยราช

ภัฏมหาสารคาม จำนวน 111 โรงเรียน ซึ่งแต่ละโรงเรียนมีบริบทที่แตกต่างกัน ทั้งด้านการเรียนการสอน เครื่องมือ ความห่างไกล จำนวนครูผู้สอน และนักเรียน

2.4 การเรียนรู้สื่ออีดีแอลทีวี เป็นกิจกรรมส่งเสริมครูให้มีการประยุกต์ใช้สื่ออีดีแอลทีวีในการเรียนการสอนตามบริบทของโรงเรียน ทั้งแบบออนไลน์และแบบออฟไลน์ ประกอบด้วย 4 กลุ่ม คือ

- 1) การใช้งานออนไลน์ผ่านเว็บไซต์
- 2) การใช้งานออนไลน์ผ่านคอมพิวเตอร์แท็บเล็ต
- 3) การใช้งานออฟไลน์แบบเซิร์ฟเวอร์
- 4) การใช้งานออฟไลน์แบบเครื่องเดียว

2.5 กิจกรรมค่ายอาสา เป็นการดำเนินงานร่วมกันของหน่วยงานทั้งภาครัฐและเอกชน ประกอบด้วย มหาวิทยาลัยราชภัฏมหาสารคาม สพป. สถานประกอบการ และชุมชน เพื่อจัดกิจกรรมค่ายอาสาเพื่อการเรียนรู้สื่ออีดีแอลทีวีตามบริบทของโรงเรียนขนาดเล็ก โดยมีการดำเนินงาน 4 ขั้นตอนคือ ขั้นวางแผน ขั้นปฏิบัติ ขั้นสังเกต และขั้นสะท้อนผล โดยแต่ละขั้นการดำเนินงานจะใช้กระบวนการแบบปัญหาเป็นฐาน ดังนี้

2.5.1 ขั้นวางแผน (Planning) เป็นการเตรียมความพร้อมของหน่วยงานความร่วมมือ เพื่อปรึกษาหารือ แนวทางการจัดกิจกรรมค่ายอาสา

2.5.2 ขั้นปฏิบัติ (Action) เป็นการจัดกิจกรรมการเรียนรู้สื่ออีดีแอล และการประยุกต์ใช้สื่ออีดีแอลทีวีเพื่อการเรียนการสอนที่โรงเรียน ตามแนวทางการดำเนินงานที่กำหนด

2.5.3 ขั้นสังเกต (Observation) เป็นการศึกษาผลการประยุกต์ใช้สื่ออีดีแอลทีวีในโรงเรียน หลังจากที่โรงเรียนและครูได้เข้าร่วมกิจกรรมค่ายอาสาตามรูปแบบที่สังเคราะห์ขึ้น

2.5.4 ขั้นสะท้อนผล (Reflection) เป็นการจัดกิจกรรมรับฟังความคิดเห็นจากผู้เข้าร่วมกิจกรรมค่ายอาสา โดยสะท้อนออกมาในรูปแบบการเขียนเรียนความเรื่องเล่า ทั้งผู้บริหาร ครู นักเรียน ชุมชน และวิทยากรของมหาวิทยาลัย

2.6 หน่วยงานความร่วมมือ เป็นหน่วยงานที่ช่วยส่งเสริม สนับสนุนการจัดกิจกรรมค่ายอาสา ประกอบด้วย 3 ส่วน ดังนี้

2.6.1 สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน สำนักงานเขตพื้นที่การศึกษาประถมศึกษาเป็นหน่วยงานที่กำหนดนโยบายการใช้สื่อไอซีทีและการใช้สื่ออีดีแอลทีวี เพื่อการเรียนการสอน สนับสนุนให้บุคลากรให้คำปรึกษา นิเทศ และติดตามประเมินผล

2.6.2 มหาวิทยาลัยราชภัฏมหาสารคาม เป็นหน่วยงานให้บริการเผยแพร่ ถ่ายทอดและพัฒนาสื่อการเรียนรู้

การสอน สนับสนุนบุคลากร วิทยากร และ เครื่องมือการจัดกิจกรรมค่ายอาสาเพื่อการเรียนรู้สื่ออีดีแอลทีวี

2.6.3 หน่วยงานสนับสนุนด้านนโยบายและงบประมาณ

- 1) สำนักงานเลขาธิการโครงการเทคโนโลยีสารสนเทศตามพระราชดำริฯ ให้การสนับสนุนสื่ออีดีแอลทีวีสำหรับโรงเรียนขนาดเล็ก
- 2) สหกรณ์ออมทรัพย์ครูให้การสนับสนุนสื่ออีดีแอลทีวีโรงเรียนขนาดเล็กในจังหวัดมหาสารคาม
- 3) ชุมชน ให้การสนับสนุนในเรื่องงบประมาณการจัดกิจกรรมเพื่อการเรียนรู้ของนักเรียนด้วยสื่ออีดีแอลทีวี
- 4) สถานประกอบการ ให้การสนับสนุนบุคลากร ให้ความรู้และส่งเสริมการใช้เทคโนโลยีสารสนเทศ

2.7 ตัวชี้วัดการดำเนินงาน

- 1) มีการวางแผนเพื่อการเรียนรู้สื่ออีดีแอลทีวีของโรงเรียนร่วมกับสมาชิกเครือข่าย
- 2) ครูสามารถประยุกต์ใช้สื่ออีดีแอลทีวีในการเรียนการสอนตามบริบทของโรงเรียน
- 3) ครูมีความรู้ ทักษะ และเจตคติต่อการประยุกต์ใช้สื่ออีดีแอลทีวีเพื่อการเรียนรู้
- 4) โรงเรียนมีเครือข่ายความร่วมมือกับชุมชนและมหาวิทยาลัยราชภัฏ
- 5) ครูมีช่องทางการติดต่อสื่อสาร การให้คำแนะนำปรึกษาการประยุกต์ใช้สื่ออีดีแอลทีวีเพื่อการเรียนการสอน

4.3 ผลการศึกษาความคิดเห็นของผู้เชี่ยวชาญที่มีต่อรูปแบบการจัดกิจกรรมค่ายอาสาเพื่อการเรียนรู้สื่ออีดีแอลทีวี

ตารางที่ 2: ความคิดเห็นผู้เชี่ยวชาญที่มีต่อรูปแบบ

รายการ	ระดับความคิดเห็น		
	X	S.D.	ความหมาย
1. ด้านกระบวนการ	4.64	0.01	มากที่สุด
1.1 การใช้ PAOR ในกิจกรรม	4.67	0.50	มากที่สุด
1.2 ขั้นวางแผน (Planning)	4.67	0.50	มากที่สุด
1.3 ขั้นปฏิบัติ (Action)	4.56	0.53	มากที่สุด
1.4 ขั้นสังเกต (Observation)	4.67	0.50	มากที่สุด
1.5 ขั้นสะท้อนผล (Reflection)	4.67	0.50	มากที่สุด
2. ด้านเทคนิคการเรียนรู้	4.57	0.01	มากที่สุด
2.1 การเรียนรู้แบบปัญหาเป็นฐาน	4.56	0.53	มากที่สุด
2.2 ความเหมาะสมของการเรียนรู้แบบปัญหาเป็นฐานกับกิจกรรมค่าย	4.45	0.53	มาก
2.3 ความเหมาะสมของกิจกรรมความร่วมมือแบบ Learning Together	4.56	0.53	มากที่สุด

ตารางที่ 2: (ต่อ)

รายการ	ระดับความคิดเห็น		
	\bar{X}	S.D.	ความหมาย
2.4 ความสอดคล้องของเป้าหมายการเรียนรู้ด้วยกิจกรรมค่ากับการเรียนรู้แบบ Learning Together	4.56	0.53	มากที่สุด
2.5 การเรียนรู้แบบ Coaching	4.56	0.53	มากที่สุด
2.6 ความสอดคล้องของเป้าหมายการเรียนรู้ด้วยกิจกรรมค่ากับการเรียนรู้แบบ Coaching	4.67	0.50	มากที่สุด
3. ด้านการสื่อสาร	4.69	0.04	มากที่สุด
3.1 การแลกเปลี่ยนเรียนรู้ผ่านเครือข่าย : Class Start	4.78	0.44	มากที่สุด
3.2 เป้าหมายการแลกเปลี่ยนเรียนรู้ผ่านเครือข่าย : Class Start	4.67	0.50	มากที่สุด
3.3 การใช้สื่ออีดีแอลทีวี แบบ online	4.67	0.50	มากที่สุด
3.4 การใช้สื่ออีดีแอลทีวีแบบ offline	4.67	0.50	มากที่สุด
4. ด้านนโยบายภาครัฐ	4.67	0.50	มากที่สุด
5. ด้านบริบทโรงเรียนขนาดเล็ก	4.67	0.50	มากที่สุด
6. ด้านหน่วยงานความร่วมมือ	4.72	0.09	มากที่สุด
6.1 สพฐ./สพป.	4.67	0.50	มากที่สุด
6.2 โรงเรียน	4.78	0.44	มากที่สุด
6.3 มหาวิทยาลัยราชภัฏมหาสารคาม	4.89	0.33	มากที่สุด
6.4 หน่วยงานสนับสนุน	4.56	0.53	มากที่สุด
7. ด้านตัวชี้วัดรูปแบบกิจกรรม	4.78	0.44	มากที่สุด
เฉลี่ยโดยรวมทุกด้าน	4.65	0.05	มากที่สุด

จากตารางที่ 2 พบว่า ความคิดเห็นของผู้เชี่ยวชาญที่มีต่อรูปแบบกิจกรรมค่าอาสาเพื่อการเรียนรู้สื่ออีดีแอลทีวีอยู่ในระดับมากที่สุด ($\bar{X} = 4.65$, S.D. = 0.05) เมื่อพิจารณารายด้าน พบว่า ทุกด้านมีความเหมาะสมมากที่สุด โดยมีค่าเฉลี่ยอยู่ระหว่าง 4.57 ถึง 4.78 และ S.D. อยู่ระหว่าง 0.01 ถึง 0.50

5. สรุปผลการวิจัย

การวิจัยในครั้งนี้ พบว่า โรงเรียนขนาดเล็กมีความต้องการเข้าร่วมกิจกรรมค่าอาสา แบบกลุ่มโรงเรียน แบบแยกโรงเรียน กิจกรรมแบบเซิร์ฟเวอร์ และกิจกรรมแบบนักศึกษาฝึกสอนตามลำดับ ทั้งนี้เนื่องจากโรงเรียนขนาดเล็กเป็นโรงเรียนที่มีบริบท ความพร้อม สภาพแวดล้อม และงบประมาณที่แตกต่างกัน จากการศึกษาความต้องการของโรงเรียนในการจัดกิจกรรมค่า พบว่า โรงเรียนต้องการเรียนรู้วิธีการใช้งานสื่ออีดีแอลทีวีแบบออนไลน์และแบบออฟไลน์ การใช้งานผ่านคอมพิวเตอร์ แท็บเล็ต การใช้งานผ่านระบบเครือข่ายภายในของโรงเรียน และการใช้งานแบบเซิร์ฟเวอร์ ตามลำดับ ความต้องการ

กิจกรรมเสริมในขณะจัดกิจกรรมค่าอาสา ได้แก่ ต้องการเรียนรู้การสำเนาสื่อ การซ่อมบำรุงระบบและเครือข่าย และการพัฒนาหนังสืออิเล็กทรอนิกส์ตามลำดับ และผลการสังเคราะห์รูปแบบการจัดกิจกรรมค่าอาสาเพื่อการเรียนรู้สื่ออีดีแอลทีวีสำหรับโรงเรียนขนาดเล็ก เครือข่ายมหาวิทยาลัยราชภัฏมหาสารคาม มีองค์ประกอบ 7 ส่วน คือ 1) นโยบายภาครัฐ 2) สื่ออีดีแอลทีวี 3) โรงเรียนขนาดเล็ก 4) การเรียนรู้สื่ออีดีแอลทีวี 5) กิจกรรมค่าอาสา 6) หน่วยงานความร่วมมือ และ 7) ตัวชี้วัด ทั้งนี้ความคิดเห็นของผู้เชี่ยวชาญที่มีต่อรูปแบบการจัดกิจกรรมค่าอาสาเพื่อการเรียนรู้สื่ออีดีแอลทีวีสำหรับโรงเรียนขนาดเล็ก ที่ผู้วิจัยสังเคราะห์ขึ้น มีความเหมาะสมในระดับมากที่สุด

เอกสารอ้างอิง

- [1] กิตติยา เศษวรรณวุฒิ. กิจกรรมค่าอาสา. <http://www.gotoknow.org/posts/258414>. [7 พ.ค.56]
- [2] คณะเทคโนโลยีสารสนเทศ. รายงานสรุปผลการจัดกิจกรรมค่าอาสา 2553-2554 มหาวิทยาลัยราชภัฏมหาสารคาม, 2554.
- [3] โครงการจัดทำเนื้อหาาระบบ e-Learning ของการศึกษาทางไกลผ่านดาวเทียม. <http://edltv.thai.net>. [20 เมษายน 2556]
- [4] บุญชม ศรีสะอาด. การวิจัยเบื้องต้น. กรุงเทพฯ, 2545.
- [5] มหาวิทยาลัยราชภัฏมหาสารคาม. รายงานผลการดำเนินงานการส่งเสริม เผยแพร่และพัฒนาต่อยอดสื่อ eDLTV. มหาวิทยาลัยราชภัฏมหาสารคาม, 2554.
- [6] มหาวิทยาลัยราชภัฏมหาสารคาม. รายงานกิจกรรมเรื่องเล่าโครงการค่าอาสาเยาวชนทั่วถิ่นไทย เรียนรู้ได้พร้อมพระบารมี. มหาวิทยาลัยราชภัฏมหาสารคาม, 2556.
- [7] วรปภา อารีราษฎร์ และคณะ. รูปแบบกิจกรรมค่าอาสาเครือข่ายมหาวิทยาลัยราชภัฏมหาสารคาม. มหาวิทยาลัยราชภัฏมหาสารคาม, 2556.
- [8] สมเจตน์ ภูศรี และคณะ. รูปแบบการเผยแพร่ ถ่ายทอดและพัฒนาสื่อ eDLTV เครือข่ายความร่วมมือมหาวิทยาลัยราชภัฏมหาสารคาม. รายงานการดำเนินงานเครือข่ายเผยแพร่ ถ่ายทอด และพัฒนาสื่อการเรียนการสอนบนระบบ e-Learning (eDL-Square). มหาวิทยาลัยราชภัฏมหาสารคาม, 2555.

การศึกษานวัตกรรมสื่ออีดีแอลทีวีเพื่อการเรียนการสอน

A Study on the Innovation in eDLTV Media for Instruction

วโรปภา อารีราษฎร์(Worapapha Arreerard)¹ ละองทิพย์ มัทธูรส (Laongtip Mathurasa)²

มนต์ชัย เทียนทอง (Monchai Tiantong)³ และ ดุษณี สุภวรรธนกุล (Dusanee Supawantanakul)²

¹นักศึกษาคณะศึกษาศาสตร์ มหาวิทยาลัยราชภัฏพระนคร ²คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏพระนคร

³คณะครุศาสตร์อุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹dr.worapapha@hotmail.com, ²laongtip77@hotmail.com, ³drmonchai@gmail.com,s.dusanee@hotmail.com

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ 1) ศึกษาหลักการและแนวคิดของนวัตกรรมที่ใช้ในการวิเคราะห์สื่ออีดีแอลทีวี และ 2) ศึกษาความคิดเห็นของผู้เชี่ยวชาญที่มีต่อผลการวิเคราะห์สื่ออีดีแอลทีวีตามหลักการแนวคิดของนวัตกรรม กลุ่มเป้าหมายเป็นผู้เชี่ยวชาญ 3 กลุ่ม กลุ่มละ 3 คน รวมทั้งสิ้น 9 คน คือด้านการจัดการเทคโนโลยี ด้านเทคโนโลยีสารสนเทศ และด้านคอมพิวเตอร์ศึกษา เครื่องมือและเทคนิคที่ใช้คือ การระดมสมองและแบบสอบถามความคิดเห็นของผู้เชี่ยวชาญที่มีต่อผลการวิเคราะห์สื่ออีดีแอลทีวีตามหลักการแนวคิดของนวัตกรรม และสถิติที่ใช้คือ ร้อยละ

ผลการวิจัยพบว่า 1) หลักการและแนวคิดของนวัตกรรมที่ใช้ในการวิเคราะห์สื่ออีดีแอลทีวี ประกอบด้วย 7 ประเด็น คือ 1.1) ประเภทและความสำคัญของเทคโนโลยีที่นำมาพัฒนาสื่ออีดีแอลทีวี 1.2) รูปแบบและประเภทของนวัตกรรมสื่ออีดีแอลทีวี 1.3) วงจรชีวิตของเทคโนโลยีและนวัตกรรมสื่ออีดีแอลทีวี 1.4) ระดับความใหม่ของนวัตกรรมสื่ออีดีแอลทีวี 1.5) การเผยแพร่และการยอมรับนวัตกรรมสื่ออีดีแอลทีวี 1.6) แนวทางการพัฒนาต่อยอดนวัตกรรมสื่ออีดีแอลทีวี และ 1.7) ความสำคัญของนวัตกรรมสื่ออีดีแอลทีวีเพื่อการเรียนการสอน และ 2) ผลการศึกษาค้นคว้าความคิดเห็นของผู้เชี่ยวชาญที่มีต่อการวิเคราะห์สื่ออีดีแอลทีวีตามหลักการแนวคิดของนวัตกรรม พบว่า ความคิดเห็นของผู้เชี่ยวชาญที่มีต่อผลการวิเคราะห์เฉลี่ยโดยรวม เห็นด้วยกับการวิเคราะห์ ร้อยละ 94.12 เมื่อพิจารณารายด้านพบว่าผู้เชี่ยวชาญเห็นด้วยกับผลการวิเคราะห์ ตามลำดับดังนี้ ด้านนวัตกรรมสื่ออีดีแอลทีวีได้รับการเผยแพร่และการยอมรับจากโรงเรียน ครูสถานศึกษา อย่างเป็นทางการและเป็นรูปธรรม และด้านความสำคัญและมีประโยชน์ในการนำไปใช้เพื่อการเรียนการสอนของโรงเรียนและสถานศึกษา ร้อยละ 100 ด้านแนวทางการพัฒนาต่อยอดนวัตกรรมสื่ออีดีแอลทีวีในบริบทของมหาวิทยาลัยราชภัฏ สามารถทำได้ 3 แนวทาง ร้อยละ 96.30 ด้านรูปแบบและประเภทของนวัตกรรมสื่ออีดีแอลทีวี ร้อยละ 94.44 ด้านประเภทและความสำคัญของ

เทคโนโลยีที่นำมาพัฒนา และวงจรชีวิตของเทคโนโลยีและนวัตกรรมสื่ออีดีแอลทีวี ร้อยละ 92.59 และด้านระดับความใหม่ของนวัตกรรมสื่ออีดีแอลทีวี ร้อยละ 91.67

คำสำคัญ: สื่ออีดีแอลทีวี นวัตกรรม เทคโนโลยี สื่อการเรียนการสอน

Abstract

The objectives of the research were 1) to study the principles and concepts of innovation for the analysis of eDLTV media and 2) to assess the results of eDLTV media analysis by experts. The target population were nine experts in technology management, information technology and computer education. The research using Brainstorming technique data analysis and the instrument was an assessment form of eDLTV media analysis. The statistics used was percentage.

The research finding showed that: 1) The principles and concepts of eDLTV analysis consisted of seven main points : types and importance of technology for eDLTV development, forms and types of eDLTV Innovation, Life circle of technology and eDLTV Innovation, modernity of eDLTV Innovation, diffusion and adoption of eDLTV Innovation and further study on eDLTV development and importance of eDLTV Innovation for instruction. 2) Regarding the results assessed by nine experts, the finding indicated that the average percentage of eDLTV media analysis was 94.12%. The percentage of diffusion and adoption of eDLTV media and importance of eDLTV media for instruction were 100%, The percentage of further study on eDLTV development was 96.30%, The percentage of forms and types of eDLTV Innovation was 94.44%, The percentage of types and importance of the technology for eDLTV development and

the percentage of lifecycle were 92.59 %, and the percentage of modernization of eDLTV Innovation was 91.67%.

Keywords: eDLTV Media, Innovation, Technology, Instructional Media

1. บทนำ

สำนักงานนวัตกรรมแห่งชาติ [7] ได้ให้นิยามของ นวัตกรรม คือ สิ่งใหม่ที่เกิดจากการใช้ความรู้และความคิดสร้างสรรค์ ที่มีประโยชน์ต่อเศรษฐกิจและสังคม สื่ออีดีแอลทีวี [4] เป็นสื่อที่จัดทำขึ้นโดยโครงการจัดทำเนื้อหาบนระบบ e-Learning ของการศึกษาทางไกลผ่านดาวเทียม เฉลิมพระเกียรติ เนื่องในโอกาสมหามงคลเฉลิมพระชนมพรรษา 80 พรรษา 5 ธันวาคม 2550 หรือเรียกว่า “สื่ออีดีแอลทีวี” มาจากคำว่า Electronic Distance Learning Via Television : eDLTV จัดทำขึ้นสืบเนื่องมาจากโรงเรียนภายใต้โครงการเทคโนโลยีสารสนเทศตามพระราชดำริ สมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี หรือโรงเรียนในโครงการ ทสรช. จำนวน 85 แห่ง ส่วนใหญ่เป็นโรงเรียนต่างจังหวัดอยู่ในชนบทที่ห่างไกล ประสบปัญหาการขาดแคลนครูเป็นจำนวนมาก จัดการเรียนการสอน โดยนำระบบการศึกษาทางไกลผ่านดาวเทียมมาใช้ จากการตรวจเยี่ยม โรงเรียนประจำปีของคณะกรรมการโรงเรียน ทสรช. พบว่า [4] โรงเรียนในโครงการ ทสรช. ยังคงประสบปัญหาการจัดการเรียนการสอนโดยใช้ระบบการศึกษาทางไกลผ่านดาวเทียม อาทิ นักเรียนจดบันทึก หรือเรียนไม่ทันการสอนของโรงเรียนวังไกลกังวล ครูต้องการสื่อประกอบการสอนเพื่อทบทวนให้แก่ นักเรียน และในกรณีที่ โรงเรียนมีปัญหาด้านครูไม่ครบชั้น การขาดแคลนครู ทำให้ครูต้องสอนหลายวิชา ปัญหาขาดแคลนสื่อการสอน ครูที่มีวุฒิการศึกษาไม่ตรงสาขาวิชา ปัญหาการฝึกทักษะการสอนในวิชาที่ครูไม่ชำนาญ หรือ ไม่มีความพร้อมเรื่องเครื่องมือ และอุปกรณ์ระบบการศึกษาทางไกลผ่านดาวเทียม จึงมีความจำเป็นอย่างยิ่ง

โครงการเทคโนโลยีสารสนเทศตามพระราชดำริ สมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี โดยศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ หรือ NECTEC สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ หรือ สวทช. กระทรวงวิทยาศาสตร์และเทคโนโลยี ร่วมกับ มูลนิธิการศึกษาทางไกลผ่านดาวเทียม นำเนื้อหาที่ถ่ายทอดมาจัดทำเป็นสื่ออิเล็กทรอนิกส์ในรูปแบบ e-Learning ภายใต้ระบบการจัดการเรียนการสอน eDL-Square Version 1 ที่พัฒนาโดย NECTEC เข้าถึงได้ทั้งแบบออนไลน์ (On-Line) ที่เว็บไซต์ www.edlvtv.thai.net และแบบออฟไลน์ (Off-Line) ที่บรรจุไว้ในฮาร์ดดิสก์เผยแพร่ โดยมหาวิทยาลัยราชภัฏ 35 แห่ง [6]

จากการที่เทคโนโลยีสารสนเทศและการสื่อสาร มีการพัฒนาอย่างต่อเนื่องและมีบทบาทต่อการเรียนรู้ของเยาวชน [2] การดำเนินงานตามภารกิจการเผยแพร่ ถ่ายทอดและพัฒนาสื่ออีดีแอลทีวีเพื่อการเรียนการสอน ของมหาวิทยาลัยราชภัฏ ในยุคเทคโนโลยีมีการพัฒนาอย่างต่อเนื่องและรวดเร็ว [9] การที่จะให้สื่ออีดีแอลทีวีเป็นสื่อที่มีคุณค่าและเป็นสื่อทางเลือกของครูนำไปใช้ในการเรียนการสอน จำเป็นต้องหากระบวนการให้ครูได้รู้จักสื่อ เข้าใจ เข้าถึงและเข้าใช้สื่ออีดีแอลทีวีในการเรียนการสอน ซึ่งจะทำให้ครูได้เห็นประโยชน์และคุณค่าของสื่ออีดีแอลทีวีต่อการเรียนรู้ของนักเรียน ผู้วิจัยจึงศึกษา สื่ออีดีแอลทีวีตามแนวคิดและหลักการของนวัตกรรม เพื่อให้ทราบถึงสถานะของสื่ออีดีแอลทีวี หรือ นวัตกรรมสื่ออีดีแอลทีวีที่เป็นอยู่ปัจจุบัน ผลจากการศึกษาจะนำมาเป็นแนวทางในการพัฒนานวัตกรรม เพื่อให้บริการสื่ออีดีแอลทีวีสอดคล้องกับความต้องการของครู ในบริบทของมหาวิทยาลัยราชภัฏ เพื่อให้สื่ออีดีแอลทีวีเป็นสื่อที่มีคุณค่าและมีประโยชน์ต่อครูในการนำไปใช้เพื่อการเรียนการสอนตรงตามวัตถุประสงค์และบรรลุเป้าหมายของโครงการตามพระราชดำริฯ ที่ส่งเสริมให้สื่ออีดีแอลทีวีเป็นสื่อที่ “เยาวชนทั่วถิ่นไทย เรียนรู้ได้ ไร้พรมพรม” สืบต่อไป

2. วัตถุประสงค์การวิจัย

2.1 เพื่อศึกษาหลักการและแนวคิดของนวัตกรรมที่ใช้ในการวิเคราะห์สื่ออีดีแอลทีวี

2.2 เพื่อศึกษาความคิดเห็นของผู้เชี่ยวชาญที่มีต่อผลการวิเคราะห์สื่ออีดีแอลทีวีตามหลักการแนวคิดของนวัตกรรม

3. วรรณกรรมที่เกี่ยวข้อง

ในช่วงที่ผ่านมาวิทยาศาสตร์และเทคโนโลยี มีบทบาทสำคัญเพิ่มขึ้น [3] จนสามารถสร้าง นวัตกรรม (Innovation) คือ การเรียนรู้ การผลิต และการใช้ประโยชน์จากความคิดใหม่ ให้เกิดผลทั้งทางเศรษฐกิจ สังคม การเมือง สิ่งแวดล้อมและวัฒนธรรม จากความก้าวหน้าของเทคโนโลยีสารสนเทศ ช่วยให้เกิดการเปลี่ยนแปลงในการทำงาน โดยเพิ่มประสิทธิภาพและประสิทธิผลของงานยิ่งขึ้น การนำเทคโนโลยีมาใช้กับงาน มีส่วนช่วยในด้านประสิทธิภาพ (Efficiency) ประสิทธิภาพ (Productivity) และประหยัด (Economy) [1] หากพิจารณาเทคโนโลยีและนวัตกรรมขององค์กร จะพิจารณาการเกิดขึ้นของเทคโนโลยี จากปัจจัยดังต่อไปนี้

3.1 ประเภทและความสำคัญของเทคโนโลยี

3.1.1 ระยะเวลาหรือช่วงของเทคโนโลยี มีดังนี้ [8]

1) เทคโนโลยีแรกเกิด (Emerging Technology) เป็นเทคโนโลยีเกิดใหม่ ยังใช้งานอยู่ในวงจำกัดในปัจจุบัน แต่คาดว่าจะถูกนำไปใช้ในเชิงพาณิชย์อย่างเต็มที่ในอนาคตอันใกล้

2) เทคโนโลยีใหม่(New Technology) เป็นเทคโนโลยีที่เพิ่งนำมาใช้ในองค์กร และส่งผลกระทบต่ออย่างชัดเจนต่อการดำเนินงานขององค์กร

3) เทคโนโลยีขั้นสูง(High Technology) เป็นเทคโนโลยีที่ซับซ้อน มีการเปลี่ยนแปลงรวดเร็วและจะมีความรู้ทางทฤษฎีเป็นองค์ประกอบที่ชัดเจน

4) เทคโนโลยีขั้นกลาง(Medium Technology) เป็นเทคโนโลยีที่อยู่ระหว่างขั้นต่ำและขั้นสูง โดยมากมักจะเป็นเทคโนโลยีที่อึดตัว และง่ายต่อการถ่ายโอน

5) เทคโนโลยีขั้นต่ำ(Low Technology) เป็นเทคโนโลยีที่ได้แพร่หลายไปสู่สังคมส่วนใหญ่ เทคโนโลยีมีความอยู่ตัวและมีการเปลี่ยนแปลงน้อยมาก

3.1.2 ระดับความสำคัญของเทคโนโลยีภายในองค์กร แบ่งออกเป็น 3 ระดับ ดังนี้

1) เทคโนโลยีโดดเด่น (Distinctive Technology) เป็นเทคโนโลยีที่ช่วยสร้างความได้เปรียบเหนือคู่แข่ง มีความแตกต่างจากคู่แข่ง สามารถนำไปใช้ประโยชน์ในเชิงพาณิชย์

2) เทคโนโลยีพื้นฐาน (Basic Technology) เป็นเทคโนโลยีที่มีความจำเป็นต่อภารกิจหลักขององค์กร แต่ไม่ทำให้องค์กรโดดเด่น

3) เทคโนโลยีภายนอก (External Technology) เป็นเทคโนโลยีที่ได้มาจากหน่วยงานภายนอก หาได้ง่ายในท้องตลาด และไม่มีผลสำคัญต่อความอยู่รอดขององค์กร

จากความสำคัญของเทคโนโลยีที่มีผลต่อองค์กร การพิจารณาใช้เทคโนโลยีที่เหมาะสม (Appropriate Technology) กับทรัพยากรที่มีอยู่ในองค์กร จะมีบทบาทสำคัญในการถ่ายทอดเทคโนโลยี ทำให้ใช้ทรัพยากรได้เต็มประสิทธิภาพและนำไปสู่การผลิตที่สูงกว่า

3.2 รูปแบบและประเภทของนวัตกรรม

นวัตกรรมที่พัฒนาขึ้นมีรูปแบบตามวัตถุประสงค์ของการใช้งานและมีประเภทที่แตกต่างกัน ดังนี้ [2]

3.2.1 รูปแบบของนวัตกรรม สำนักงานนวัตกรรมแห่งชาติ ได้จำแนกเป็น 3 รูปแบบ คือ นวัตกรรมผลิตภัณฑ์ (Product Innovation) นวัตกรรมบริการ (Service Innovation) และ นวัตกรรมกระบวนการ (Process Innovation)

3.2.2 ประเภทของนวัตกรรมจำแนกตามระดับความใหม่ขององค์ประกอบ (Component) หรือระบบ (System) ได้ดังนี้

1) นวัตกรรมแบบค่อยเป็นค่อยไป (Incremental Innovation) เป็นนวัตกรรมที่มีการปรับปรุงองค์ประกอบต่าง ๆ ให้ดีขึ้น ตามลำดับ และปรับปรุงอย่างต่อเนื่อง

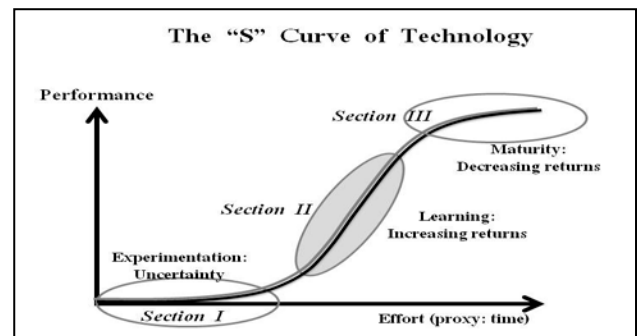
2) นวัตกรรมแบบหน่วยแยกที่สามารถรวมกันได้ (Modular Innovation) เป็นนวัตกรรมที่นำเอาองค์ประกอบใหม่มาใช้ โดยไม่มีการเปลี่ยนแปลงระบบการทำงาน มีการเปลี่ยนแปลงองค์ประกอบที่สำคัญบางส่วน โครงสร้างโดยรวมยังคงเดิม

3) นวัตกรรมสถาปัตยกรรม (Architectural Innovation) เป็นนวัตกรรมที่มีการปรับโครงสร้าง ของระบบที่มีอยู่เดิมเพื่อเชื่อมโยงองค์ประกอบต่าง ๆ ในรูปแบบใหม่

4) นวัตกรรมแบบก้าวกระโดด (Radical Innovation) เป็นนวัตกรรมที่มีความใหม่ที่แตกต่างไปจากกรรมวิธีและแนวคิดเดิมอย่างเห็นได้ชัดเจน

3.3 วงจรชีวิตของเทคโนโลยีและนวัตกรรม

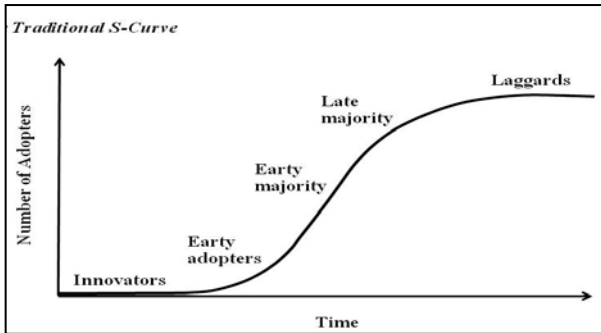
Rogers [11] ได้สร้างวัฏจักรการเปลี่ยนแปลงเทคโนโลยี หรือ S-Curve อธิบายกระบวนการแพร่กระจายของเทคโนโลยีในสังคมเพื่อที่จะคาดการณ์ช่วงเวลา การยอมรับเทคโนโลยีและกระบวนการแพร่กระจายเทคโนโลยีในสังคม ดังภาพที่ 1



ภาพที่ 1: วงจรชีวิตของเทคโนโลยี

ที่มา : Rogers, 1962 : 32

จากภาพที่ 1 แกน Y แทนประสิทธิภาพหรือเทียบจำนวนผู้ใช้ในสังคม แกน X เป็นเวลา Section I เป็นช่วงเวลาของการประดิษฐ์คิดค้นเทคโนโลยีหรือนวัตกรรม Section II เป็นช่วงเวลาที่นวัตกรรมมีการปฏิสัมพันธ์กับคนในสังคม Section III เป็นช่วงเวลาที่นวัตกรรมอึดตัว การพัฒนานวัตกรรมอาจจะไม่สามารถพัฒนาต่อไปได้ ทั้งนี้ Rogers ได้อธิบาย การเกิด การนิยม และการตกต่ำของนวัตกรรม หรือ “หุบเหว” แห่งนวัตกรรม ซึ่งเกิดจากการยอมรับนวัตกรรมของกลุ่มคนในสังคม ดังภาพที่ 2



ภาพที่ 2: การยอมรับนวัตกรรมของกลุ่มคน
ที่มา : Rogers, 1962 : 32

จากภาพที่ 2 การยอมรับนวัตกรรมของกลุ่มคนในสังคมจะมีความแตกต่างกัน ตามลักษณะเฉพาะของแต่ละกลุ่ม ดังนี้ [12]

กลุ่ม Inventor คือ คนกลุ่มแรกในสังคม ที่เป็นทั้งผู้ประดิษฐ์คิดค้น เป็นผู้ใช้งาน และชอบติดตามเทคโนโลยีอยู่เสมอ

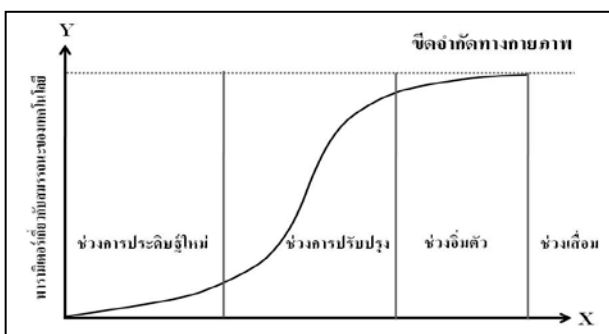
กลุ่ม Early Adopters เป็นกลุ่มที่ชอบลองอะไรใหม่ๆ และค่อนข้างมีฐานะ อาจเป็นนักวิชาการหรือคนมีชื่อเสียงในสังคม

กลุ่ม Early Majority คนกลุ่มนี้การตัดสินใจเลือกนวัตกรรมจะต้องคิดหลายรอบ และมักดูจากการตัดสินใจของสองกลุ่มแรก

กลุ่ม Late Majority คนกลุ่มนี้จะใช้เทคโนโลยีหรือนวัตกรรมเมื่อมีความจำเป็นต้องการใช้งานจริงๆ เท่านั้น

กลุ่ม Laggard เป็นกลุ่มที่ใช้เทคโนโลยีหรือนวัตกรรมเมื่อตก runt ไปแล้ว และการเลือกซื้อจะดูพฤติกรรมของคนในกลุ่มอื่นก่อน

มณฑล ศาสนนันท์ [8] ได้แบ่งเส้นโค้งวงจรชีวิตของช่วงเทคโนโลยี ดังภาพที่ 3



ภาพที่ 3: เส้นโค้งวงจรชีวิตของเทคโนโลยี
ที่มา : มณฑล ศาสนนันท์, 2552 : 11

จากภาพที่ 3 เส้นโค้งวงจรชีวิตของเทคโนโลยี มี 4 ช่วง ดังนี้
ช่วงการประดิษฐ์ใหม่ เป็นช่วงของการออกแบบ การทดลอง หรือ การแก้ไขข้อผิดพลาดเบื้องต้น

ช่วงการปรับปรุง เป็นช่วงที่เทคโนโลยีเข้าสู่ท้องตลาด และจะมีการปรับปรุงสมรรถนะอย่างรวดเร็ว

ช่วงอิมิตัว เทคโนโลยีเริ่มถึงขีดจำกัดด้านสมรรถนะ ทำให้มีโอกาสถูกทดแทนหรือล้าสมัย เมื่อมีเทคโนโลยีใหม่เกิดขึ้น

ช่วงเสื่อม เป็นช่วงที่เทคโนโลยีล้าสมัย และอาจถูกทดแทนด้วยเทคโนโลยีอื่น

3.4 ระดับความใหม่ของนวัตกรรม

ระดับความใหม่ของนวัตกรรม (Newness of Innovation) British Standard [13] ได้จัดความใหม่ของนวัตกรรม เป็น 8 ระดับ เริ่มจากระดับความใหม่สำหรับตนเอง (New to an Individual) ใหม่สำหรับแผนก (New to a Department) ซึ่งทั้งสองระดับจะเป็นการใช้เทคโนโลยีที่มีอยู่แล้ว นำมาพัฒนาใช้ในหลักการที่แตกต่างไปจากเดิม ความสำคัญของการเกิดขึ้นหรือสิ่งที่ได้รับกลับมา นั้นจะมีข้อจำกัดและความเสี่ยงค่อนข้างต่ำ ส่วนความใหม่สำหรับส่วน (New to a Site) ใหม่สำหรับองค์กร (New to Company) ใหม่สำหรับตลาด (New to a Market) ใหม่สำหรับอุตสาหกรรม (New to an Industry) ใหม่สำหรับประเทศ (New to a Country) และใหม่สำหรับโลก (New to the World) จะมีการเปลี่ยนแปลงที่เห็น ได้ชัด สามารถนำเข้าสู่ในเชิงพาณิชย์ ทั้งนี้ความใหม่สำหรับประเทศ หรือความใหม่สำหรับโลก จะเป็นความใหม่ระดับสมบูรณ์ คือ เป็นความใหม่ที่ไม่เคยเห็น และไม่เคยมีมาก่อน สิ่งที่ได้จะมีความสำคัญมากและได้รับผลตอบแทนค่อนข้างสูง รวมถึงความเสี่ยงที่เกิดขึ้นก็จะสูงตามด้วย

3.5 การเผยแพร่ นวัตกรรม

Everett M Rogers and Shoemaker F. [12] ได้ให้ความหมายการเผยแพร่ (Diffusion) คือ กระบวนการซึ่งนวัตกรรมถูกสื่อสารผ่านช่องทางในช่วงเวลาหนึ่งระหว่างสมาชิกต่างๆ ที่อยู่ในระบบสังคม โดยมีปัจจัยที่เป็นส่วนประกอบของการเผยแพร่ นวัตกรรม 4 ประการ ได้แก่ ตัวนวัตกรรม (Innovation) ช่องทางการสื่อสาร (Communication Channels) เวลา (Time) และ ระบบสังคม (Social System) ที่มีธรรมชาติ วัฒนธรรม ฐานะทางเศรษฐกิจ การเมือง การปกครอง มีอำนาจต่อการยอมรับนวัตกรรมเป็นอย่างมาก

3.6 การยอมรับนวัตกรรม

การยอมรับนวัตกรรม (Adoption) อรอนงค์ โรจนวิฒนบุญย์ [11] กล่าวว่า นวัตกรรมที่ยอมรับได้ง่าย โดยมากจะมีลักษณะสำคัญ 5 ประการ คือ ผลประโยชน์ที่ได้รับจากนวัตกรรม (Relation Advantage) การเข้ากันได้ดีกับสิ่งเดิม (Compatibility) ความซับซ้อน (Complexity) การทดลองได้ (Trial ability) และการสังเกตได้ (Observability) หากนวัตกรรมมีองค์ประกอบเหล่านี้ จะส่งผลให้มีการยอมรับนวัตกรรมได้ง่ายขึ้น

ผู้วิจัยได้นำหลักการ แนวคิด และทฤษฎี จากวรรณกรรมที่ได้ศึกษา มาเป็นแนวทางในการศึกษา วิเคราะห์ สื่ออีดีแอลทีวีเพื่อการเรียนการสอน ในการวิจัยในครั้งนี้

4. การดำเนินการวิจัย

4.1 ประชากรและกลุ่มตัวอย่าง

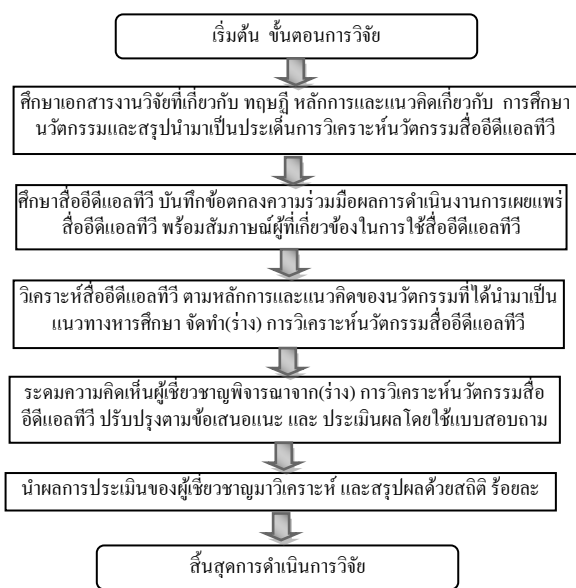
ประชากรเป็นผู้เชี่ยวชาญที่มีคุณวุฒิการศึกษาระดับปริญญาเอก คัดเลือกเป็นกลุ่มตัวอย่างแบบเจาะจง จำนวน 9 คน แบ่งเป็น 3 กลุ่ม ๆ ละ 3 คน คือ กลุ่มที่มีคุณวุฒิและประสบการณ์ด้านการจัดการเทคโนโลยี ด้านเทคโนโลยีสารสนเทศ และด้านคอมพิวเตอร์ศึกษา

4.2 เทคนิคและเครื่องมือที่ใช้ในการวิจัย

เทคนิคที่ใช้ในการวิจัยคือ เทคนิคการระดมสมอง (Brainstorming) และเครื่องมือที่ใช้ คือ แบบสอบถามความคิดเห็นของผู้เชี่ยวชาญที่มีต่อผลการวิเคราะห์สื่ออีดีแอลทีวีตามหลักการแนวคิดของนวัตกรรม มีค่าความเชื่อมั่น 0.87

4.3 สถิติที่ใช้ในการวิจัย คือ ร้อยละ

4.4 ขั้นตอนการวิจัย ดำเนินการตามลำดับ ดังภาพที่ 4 ดังนี้



ภาพที่ 4: ขั้นตอนดำเนินการวิจัย

5. ผลการดำเนินการวิจัย

5.1 ผลการศึกษาหลักการและแนวคิดของนวัตกรรมที่ใช้ในการวิเคราะห์สื่ออีดีแอลทีวี

หลักการและแนวคิดของนวัตกรรมที่ใช้ในการวิเคราะห์สื่ออีดีแอลทีวี ประกอบด้วย 7 ประเด็น คือ 1) ประเภทและความสำคัญของเทคโนโลยีที่นำมาพัฒนาสื่ออีดีแอลทีวี 2) รูปแบบและประเภทของนวัตกรรมสื่ออีดีแอลทีวี 3) วงจรชีวิตของเทคโนโลยีและนวัตกรรมสื่ออีดีแอลทีวี 4) ระดับความใหม่ของ

นวัตกรรมสื่ออีดีแอลทีวี 5) การเผยแพร่และการยอมรับ นวัตกรรมสื่ออีดีแอลทีวี 6) แนวทางการพัฒนาต่อยอด นวัตกรรมสื่ออีดีแอลทีวี และ 7) ความสำคัญของนวัตกรรมสื่ออีดีแอลทีวีเพื่อการเรียนการสอน

5.2 ผลการศึกษาความคิดเห็นของผู้เชี่ยวชาญที่มีต่อผลการวิเคราะห์สื่ออีดีแอลทีวีตามหลักการแนวคิดของนวัตกรรม การวิจัยได้ดำเนินการ 2 ขั้นตอน ดังนี้

5.2.1 ผลการสังเคราะห์สื่ออีดีแอลทีวีตามหลักการและแนวคิดของนวัตกรรม ที่ได้ศึกษาในข้อ 5.1 พบว่า

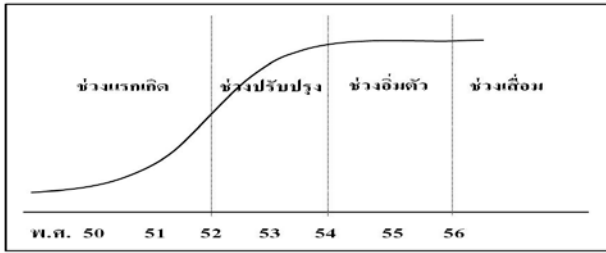
โครงการเทคโนโลยีสารสนเทศตามพระราชดำริ สมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี โดย เนคเทค และ สวทช. ร่วมกับ มูลนิธิการศึกษาทางไกลผ่านดาวเทียม จัดทำเป็นสื่ออิเล็กทรอนิกส์ในรูปแบบ e-Learning ภายใต้ระบบการจัดการเรียนการสอน eDL-Square Version 1 ที่พัฒนาต่อยอดจากระบบ Learn Square โดยใช้เทคโนโลยีโปรแกรมระบบเปิด (Open Source) เพื่อให้ได้ระบบที่สอดคล้องกับการนำเสนอเนื้อหาของโรงเรียนวังไกลกังวล ในปี พ.ศ. 2550 [5] ดังภาพที่ 5



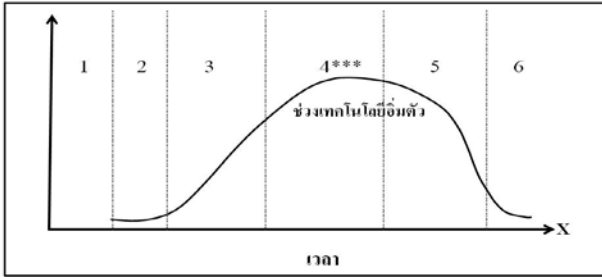
ภาพที่ 5: สถาปัตยกรรมระบบ LMS : eDL-Square V.1

หลังจากนั้นได้นำเนื้อหาการเรียนการสอนของโรงเรียนวังไกลกังวล มาพัฒนาเป็นสื่ออิเล็กทรอนิกส์ [4] นำเสนอผ่านระบบอีดีแอลสแควร์ เรียกว่า “สื่ออีดีแอลทีวี” ดังภาพที่ 6(ก) ปัจจุบันสื่ออีดีแอลทีวีมี 4 ระบบ มีสื่ออิเล็กทรอนิกส์ทั้งสิ้น 73,867 ชิ้น ครอบคลุมทุกระดับชั้น ส่งผลให้วงจรชีวิตของสื่ออีดีแอลทีวีอยู่ในช่วงเทคโนโลยีอิมตัว ดังภาพที่ 6(ข)

- ระยะที่ 1 พ.ศ. 2550 พัฒนาอีดีแอลทีวีระดับมัธยมศึกษา
- ระยะที่ 2 พ.ศ. 2553 พัฒนาอีดีแอลทีวีเพื่อพัฒนาอาชีพ
- ระยะที่ 3 พ.ศ. 2553-2554 พัฒนาอีดีแอลอาร์ยูปฐมวัย
- ระยะที่ 4 พ.ศ. 2554 พัฒนาอีดีแอลทีวีระดับประถมศึกษา



(ก) การพัฒนาสื่ออีดีแอลทีวี



(ข) เส้นโค้งวงจรชีวิตของนวัตกรรมสื่ออีดีแอลทีวี

ภาพที่ 6: วงจรชีวิตของนวัตกรรมสื่ออีดีแอลทีวี

5.2.2 ผลการประเมินความคิดเห็นของผู้เชี่ยวชาญที่มีต่อผลการวิเคราะห์สื่ออีดีแอลทีวีตามหลักการแนวคิดของนวัตกรรมนำเสนอผล ดังตารางที่ 1

ตารางที่ 1: ความคิดเห็นของผู้เชี่ยวชาญที่มีต่อผลการวิเคราะห์

รายการ	ร้อยละ
1. ประเภทและความสำคัญของเทคโนโลยีที่นำมาพัฒนา	92.59
1.1 เทคโนโลยีที่นำมาพัฒนาเป็นเทคโนโลยีขั้นกลางส่งผลให้ระบบ eDL-Square Version 1 เป็นเทคโนโลยีที่อิมตัว	88.89
1.2 สื่ออีดีแอลทีวีเป็นเทคโนโลยีพื้นฐานเป็นสื่อทางเลือกของครูและผู้สนใจทั่วไป	88.89
1.3 การพัฒนาสื่ออีดีแอลทีวีเป็นการเลือกใช้เทคโนโลยีที่เหมาะสม	100.00
2. รูปแบบและประเภทของนวัตกรรมสื่ออีดีแอลทีวี	94.44
2.1 เป็นนวัตกรรมผลิตภัณฑ์ โดยระบบ eDL-Square เป็นผลิตภัณฑ์จับต้องไม่ได้ และสื่ออีดีแอลทีวีเป็นผลิตภัณฑ์จับต้องได้	88.89
2.2 ระดับความใหม่ขององค์ประกอบหรือระบบ สื่ออีดีแอลทีวี เป็นนวัตกรรมที่มีการพัฒนาแบบค่อยเป็นค่อยไป	100.00
3. วงจรชีวิตของเทคโนโลยีและนวัตกรรมสื่ออีดีแอลทีวี	92.59
3.1 เป็นช่วงนวัตกรรม/เทคโนโลยีอิมตัว	88.89
3.2 มีการเผยแพร่ผ่านกลุ่มคนที่อยู่ระหว่างกลุ่ม Early Majority และกลุ่ม Laggard	100.00
3.3 เป็นนวัตกรรมอิมตัว การพัฒนาต่อยอดควรเป็นการพัฒนาการให้บริการที่สอดคล้องกับความต้องการของลูกค้า	88.89
4. ระดับความใหม่ของนวัตกรรมสื่ออีดีแอลทีวี	91.67
4.1 อยู่ระดับความใหม่สำหรับตนเอง และใหม่สำหรับแผนก	77.78
4.2 ไม่อยู่ในกลุ่มความใหม่สำหรับส่วน ความใหม่สำหรับองค์กร ความใหม่สำหรับตลาด และความใหม่สำหรับอุตสาหกรรม	100.00
4.3 ไม่อยู่ในกลุ่มความใหม่สำหรับประเทศ และใหม่สำหรับโลก	100.00
4.4 ด้านการเผยแพร่สื่ออีดีแอลทีวีสู่ท้องถิ่น โดยการสร้างเครือข่ายร่วมกับมหาวิทยาลัยราชภัฏถือว่าเป็นความใหม่สำหรับประเทศ	88.89
5. นวัตกรรมสื่ออีดีแอลทีวีได้รับการเผยแพร่และการยอมรับจากครู โรงเรียน และ สถานศึกษา อย่างเป็นรูปธรรม	100.00
6. แนวทางการพัฒนาต่อยอดนวัตกรรมสื่ออีดีแอลทีวี ในบริบท	96.30

ขอมหาวิทยาลัยราชภัฏ สามารถทำได้ 3 แนวทาง	
6.1 การพัฒนานวัตกรรมผลิตภัณฑ์	88.89
6.2 การพัฒนานวัตกรรมบริการ	100.00
6.3 การพัฒนานวัตกรรมกระบวนการ	100.00
7. นวัตกรรมสื่ออีดีแอลทีวีมีความสำคัญและมีประโยชน์ในการนำไปใช้เพื่อการเรียนการสอนของโรงเรียนและสถานศึกษา	100.00
ความคิดเห็นเฉลี่ยโดยรวม	94.12

จากตารางที่ 1 พบว่า ผู้เชี่ยวชาญเห็นด้วยกับผลการวิเคราะห์ โดยความคิดเห็นที่มีต่อผลการวิเคราะห์นวัตกรรมสื่ออีดีแอลทีวีเพื่อการเรียนการสอน เฉลี่ยโดยรวมร้อยละ 94.12 เมื่อพิจารณาทางด้าน พบว่า ผู้เชี่ยวชาญเห็นด้วยกับผลการวิเคราะห์นวัตกรรมสื่ออีดีแอลทีวี ความคิดเห็นอยู่ระหว่างร้อยละ 77.78 ถึง ร้อยละ 100

6. สรุปผลการวิจัย

จากผลการวิเคราะห์นวัตกรรมสื่ออีดีแอลทีวีเพื่อการเรียนการสอนตามหลักการและแนวคิดของนวัตกรรม 7 ประเด็น พบว่า ผู้เชี่ยวชาญเห็นด้วยกับผลการวิเคราะห์ นวัตกรรมสื่ออีดีแอลทีวี เฉลี่ยโดยรวมร้อยละ 94.12 โดยนวัตกรรมสื่ออีดีแอลทีวีได้รับการเผยแพร่และการยอมรับจากครู โรงเรียน สถานศึกษา อย่างเป็นรูปธรรม สื่ออีดีแอลทีวีมีความสำคัญและมีประโยชน์ในการนำไปใช้เพื่อการเรียนการสอนของโรงเรียนและสถานศึกษา รูปแบบและประเภทของนวัตกรรมสื่ออีดีแอลทีวี เป็นนวัตกรรมแบบผลิตภัณฑ์ที่มีการพัฒนาแบบค่อยเป็นค่อยไป ความใหม่ของนวัตกรรมสื่ออีดีแอลทีวีเป็นความใหม่สำหรับตนเองและความใหม่สำหรับแผนก เทคโนโลยีที่นำมาพัฒนาส่งผลให้วงจรชีวิตของสื่ออีดีแอลทีวีอยู่ในช่วงอิมตัว แนวทางการพัฒนาต่อยอดสื่ออีดีแอลทีวี ในบริบทของมหาวิทยาลัยราชภัฏ ที่ได้ 3 แนวทางคือ การพัฒนานวัตกรรมบริการ นวัตกรรมกระบวนการ และ นวัตกรรมผลิตภัณฑ์ ตามลำดับ

7. เอกสารอ้างอิง

[1] ครูกิตติกร. (2556). ความหมายของเทคโนโลยี. [Online]. Available : <http://kitikorn.site90.net/42102/basis.htm#num2>. [12 ตุลาคม 2556].

[2] สุภชัย หล่อโลหการ และคณะ. (2555). การจัดการนวัตกรรมสำหรับผู้บริหาร. พิมพ์ครั้งที่ 4. สำนักงานนวัตกรรมแห่งชาติ (องค์การมหาชน) กระทรวงวิทยาศาสตร์และเทคโนโลยี. กรุงเทพมหานคร.

[3] สุภัญญา แซ่มซ้อ. (2555). แนวคิดเชิงนวัตกรรมสำหรับการบริหารสถานศึกษาในศตวรรษที่ 21. วารสารศึกษาศาสตร์ มหาวิทยาลัยนครสวรรค์. ปีที่ 14 ฉบับที่ 2 พฤษภาคม – สิงหาคม.

[4] สำนักงานโครงการ. (2557). โครงการจัดทำเนื้อหาเพื่อการเรียนการสอนระบบ e-Learning. [Online]. Available : <http://edltv.thai.net>. [25 มกราคม 2557].

[5] สำนักงาน สวทช. (2555). บันทึกข้อตกลงความร่วมมือ เครือข่ายเผยแพร่ถ่ายทอดและพัฒนาสื่อการเรียนการสอนบนระบบ e-Learning (eDL-Square) ระยะที่ 2. สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ, กระทรวงวิทยาศาสตร์และเทคโนโลยี. กรุงเทพมหานคร.

- [6] สำนักงาน สวทช. (2553). รายงานผลการดำเนินงานโครงการเครือข่าย
เผยแพร่ ถ่ายทอด และพัฒนาสื่อการเรียนการสอนบนระบบ e-Learning
(eDL-Square). สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ,
กระทรวงวิทยาศาสตร์และเทคโนโลยี. กรุงเทพมหานคร.
- [7] สำนักงานนวัตกรรมแห่งชาติ. (2557). นิยามนวัตกรรม. [Online]. Available
: [http://www.nia.or.th/index.php?section=aboutus&page=](http://www.nia.or.th/index.php?section=aboutus&page=strategy_definition)
[strategy_definition](http://www.nia.or.th/index.php?section=aboutus&page=strategy_definition). [6 มกราคม 2557].
- [8] มณฑลีสานนันทน์. (2552). การจัดการนวัตกรรมและเทคโนโลยี.
สำนักพิมพ์มหาวิทยาลัยธรรมศาสตร์. กรุงเทพมหานคร.
- [9] มหาวิทยาลัยราชภัฏมหาสารคาม. (2556). รายงานผลการดำเนินงาน
โครงการเครือข่ายเผยแพร่ ถ่ายทอด และพัฒนาสื่อการเรียนการสอนบน
ระบบ e-Learning (eDL-square).
- [10] อรอนงค์ วัฒนบุญย์. (2553). การพัฒนาตัวแบบผู้นำเชิงนวัตกรรม.
วิทยานิพนธ์หลักสูตรรัฐประศาสนศาสตรดุษฎีบัณฑิต : คณะรัฐ
ประศาสนศาสตร์ สถาบันบัณฑิตพัฒนบริหารศาสตร์.
- [11] Rogers, Everett M. (1962). Diffusion of Innovations. 2nd ed. New York:
The Free Press of Glencoe.
- [12] Rogers, Everett M. and Shoemaker, F. Floyd. (1971). Communication of
Innovations a: Cross-Cultural Approach. New York : Free Press.
- [13] British Standard. (2006). Design management systems Part 1: Guide to
managing Innovation. University of Nottingham. Reported on 23
August 2006.

การประยุกต์ใช้สื่อ eDLTV เพื่อการเรียนรู้สู่การพัฒนาแอปพลิเคชันบนแท็บเล็ต กลุ่มสาระการเรียนรู้วิทยาศาสตร์

The Application of eDLTV media for learning to develop the application on tablet in the Substance Science Group

สิทธิพร อิทธสมบัติ (Sittiporn Ittasombut)¹ และ วรปภา อารีราษฎร์ (Worapapha Arreerard)²

สาขาคอมพิวเตอร์ศึกษา คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏมหาสารคาม

¹rider.noom@hotmail.com, ²dr.worapapha@hotmail.com

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ 1) พัฒนาหลักสูตรอบรมการประยุกต์ใช้สื่ออีดีแอลทีวีเพื่อการเรียนรู้สู่การพัฒนาแอปพลิเคชันบนแท็บเล็ต กลุ่มสาระการเรียนรู้วิทยาศาสตร์ 2) ศึกษาความรู้และความเข้าใจของครูที่เข้ารับการอบรมตามหลักสูตรอบรมที่พัฒนาขึ้น 3) ศึกษาผลการพัฒนาแอปพลิเคชันของครูที่เข้ารับการอบรมตามหลักสูตรที่พัฒนาขึ้น และ 4) ศึกษาความพึงพอใจของครูที่เข้ารับการอบรมตามหลักสูตรที่พัฒนาขึ้น กลุ่มตัวอย่างคือครูผู้สอนกลุ่มสาระการเรียนรู้วิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 1 สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษามหาสารคาม เขต 1 จำนวน 30 คน เครื่องมือที่ใช้ในการวิจัย คือ คู่มือหลักสูตรอบรม แบบประเมินหลักสูตรอบรม แบบทดสอบวัดผลการเรียนรู้ของครูที่เข้ารับการอบรม และแบบประเมินแอปพลิเคชัน ของครูที่เข้ารับการอบรม และแบบสอบถามความพึงพอใจของครูที่เข้ารับการอบรมที่มีต่อหลักสูตรอบรมที่พัฒนาขึ้น สถิติที่ใช้วิเคราะห์ข้อมูล คือ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และสถิติ *t-test* (Dependent Sample)

ผลการวิจัยพบว่า 1) หลักสูตรอบรมการประยุกต์ใช้สื่ออีดีแอลทีวีเพื่อการเรียนรู้สู่การพัฒนาแอปพลิเคชันบนแท็บเล็ต กลุ่มสาระการเรียนรู้วิทยาศาสตร์ ประกอบด้วย 6 ชั้น คือ ชั้นที่ 1 แนะนำสื่ออีดีแอลทีวีและตัวอย่างแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ต ชั้นที่ 2 เลื่อนเนื้อหา ชั้นที่ 3 ปรับปรุงเนื้อหา ชั้นที่ 4 จัดทำเกมเพื่อการเรียนรู้ ชั้นที่ 5 รวบรวมเครื่องมือ และ ชั้นที่ 6 ประเมินแอปพลิเคชันที่พัฒนาขึ้น และผลการประเมินคุณภาพหลักสูตรอบรมโดยผู้เชี่ยวชาญอยู่ในระดับเหมาะสมมากที่สุด

($\bar{X} = 4.63, S.D. = 0.49$) 2) ผู้เข้าอบรมมีความรู้และเข้าใจหลังการอบรมโดยมีคะแนนหลังการอบรมสูงกว่าคะแนนก่อนการอบรมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 3) ผลการพัฒนาแอปพลิเคชันของผู้เข้ารับการอบรม พบว่า แอปพลิเคชันที่พัฒนาขึ้นมีคุณภาพอยู่ในระดับมาก ถึง มากที่สุด ($\bar{X} = 4.33 - 4.83, S.D. = 0.37 - 0.67$) และ 4) ผู้เข้ารับการอบรมมีความพึงพอใจต่อหลักสูตรอบรมที่พัฒนาขึ้นอยู่ในระดับมากที่สุด ($\bar{X} = 4.69, S.D. = 0.47$)

คำสำคัญ: การประยุกต์ใช้ สื่ออีดีแอลทีวี แอปพลิเคชัน แท็บเล็ต

Abstract

This research aims to 1) to developed syllabus of the application of eDLTV media for learning to develop the application on tablet in the Substance Science Group 2) to study the learning and understanding of the participated teachers in this training program 3) to study the performance of developed application of the participated teachers of this training program and 4) to study the satisfaction of the participated teachers in this program. 30 science teachers of the office of Maha Sarakham Primary Education Service Area 1 were selected as the sample group. The research tools were handbooks of training syllabus, the assessment of training syllabus, the test to assess the learning perception of the participated teachers, the assessment for application of participated teachers, and the satisfaction questionnaire for the participated teacher for the syllabus. The statistics used in data analysis were

Average, Standard Deviation and t-test (Dependent Sample).

The results were found that 1) The training syllabus for the application of eDLTV media for learning to develop the application on tablet in the Substance Science Group were consisted of 6 stages: 1.1) The introduction of the eDLTV media and the example of the application for learning to develop the application on tablet. 1.2) Selected the material 1.3) improved the material 1.4) provided the games for learning 1.5) collected the tools and 1.6) assessed the developed application. The result of the qualification assessment of the training syllabus by the specialists at the highest appreciated ($\bar{X} = 4.63, S.D. = 0.49$) 2) The participated teachers were acknowledged and understand after the training syllabus had higher score at 0.05 levels statistically significant. 3) The performance of developed application of the participated teachers of this training program found that the developed application has quality at high to highest level ($\bar{X} = 4.33-4.83, S.D. = 0.37-0.67$), and 4) The participated teachers have satisfaction on the developed syllabus at the highest score ($\bar{X} = 4.69, S.D. = 0.47$).

Keywords: Apply, eDLTV media, Application, Table

1. บทนำ

แผนพัฒนาเศรษฐกิจและสังคมแห่งชาติ ฉบับที่ 11 พ.ศ. 2555-2559 ให้ความสำคัญกับการพัฒนาคุณภาพของเด็กวัยเรียน ให้มีความรู้ทางวิชาการ ทักษะและสติปัญญาที่สามารถศึกษาหาความรู้และต่อยอดองค์ความรู้ได้ด้วยตนเอง รวมทั้งสามารถปรับตัวให้รู้เท่าทันกับข่าวสาร ภายใต้บริบทแห่งการเปลี่ยนแปลงทางด้านเทคโนโลยีที่รวดเร็ว จะช่วยสนับสนุนให้เกิดการพัฒนาไปสู่ระบบการเรียนรู้ตลอดชีวิต จากนโยบายเร่งด่วนของรัฐบาลที่เร่งดำเนินการจัดหาเครื่องแท็บเล็ตพีซีให้แก่เด็กเรียนระดับชั้นประถมศึกษาปีที่ 1 ในปีการศึกษา 2555 และมอบหมายให้สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน หรือ สพฐ. ดำเนินการจัดฝึกอบรมพัฒนาบุคลากร การใช้คอมพิวเตอร์พกพา หรือคอมพิวเตอร์แท็บเล็ต ส่งเสริมครูใน

การใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนการสอน โดยจัดทำโครงการประกวดและอบรมเชิงปฏิบัติการ ผลิตภัณฑ์การเรียนการสอนเพื่อใช้กับเครื่องแท็บเล็ต ซึ่งครอบคลุมรูปแบบสื่ออิเล็กทรอนิกส์ (Multimedia eBook) การ์ตูนอิเล็กทรอนิกส์ (e-Cartoon) และ แอปพลิเคชัน การเรียนการสอน (Learning Application) โดยสื่อที่ผลิตได้จัดประกวด ร่วมกับอุปกรณ์แท็บเล็ตในโครงการ OTPC (One Tablet Per Child) พร้อมทั้งจัดเตรียมเครื่องมือสร้าง สื่อการเรียนการสอนพร้อมวัสดุอุปกรณ์ใช้ในการผลิตระหว่างอบรมเชิงปฏิบัติการ OTPC [1]

การพัฒนาแอปพลิเคชัน เพื่อการเรียนรู้ ที่ สพฐ. ดำเนินการพัฒนา 5 กลุ่มสาระการเรียนรู้ กลุ่มสาระการเรียนรู้วิทยาศาสตร์ เป็นกลุ่มสาระการเรียนรู้กลุ่มหนึ่งที่ต้องการพัฒนา โดยการสร้างสื่อกระตุ้นการเรียนรู้ การสร้างสื่อนำเสนอบทเรียน การประยุกต์การเรียนรู้ด้วยเกมหรือแบบฝึกและแบบทดสอบซึ่งเหมาะสำหรับการเรียนรู้ในระดับชั้นประถมศึกษาปีที่ 1 ที่มีการเรียนรู้ในแต่ละระยะด้วยเวลาที่สั้น จะส่งผลให้ผู้เรียนสามารถใช้เวลาในการเรียนเนื้อหาสาระ ร่วมกับกิจกรรม เกม แบบฝึกและแบบทดสอบได้อย่างง่ายและสะดวกต่อการเรียนรู้ของผู้เรียนในระดับประถมศึกษา [4]

มหาวิทยาลัยราชภัฏมหาสารคาม โดยคณะเทคโนโลยีสารสนเทศ ได้ร่วมกับ สพฐ. ส่งบุคลากรเข้ารับการอบรม เพื่อเป็นวิทยากรแกนนำ และร่วมเป็นหน่วยงานจัดอบรมให้กับครูในภาคตะวันออกเฉียงเหนือ ภายใต้การดำเนินงานโครงการศูนย์ทางไกลเพื่อพัฒนาการศึกษาและพัฒนาชนบทในภาคตะวันออกเฉียงเหนือ ร่วมกับเครือข่ายที่เป็นหน่วยงานทั้งภาครัฐและเอกชนเพื่อส่งเสริมการพัฒนาครูในการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนการสอน และการประยุกต์ใช้สื่ออีดีแอลทีวีเพื่อการเรียนการสอน [2] สื่ออีดีแอลทีวี ซึ่งเป็นสื่อภายใต้โครงการจัดทำเนื้อหาบนระบบอีเลิร์นนิ่ง (eDL-Square) ภายใต้โครงการเทคโนโลยีสารสนเทศตามพระราชดำริสมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี จัดทำขึ้นเพื่อให้ครูได้ใช้เป็นสื่อเพื่อการเรียนการสอน สื่อการเรียนการสอนในรูปแบบสื่อประสม เป็นสื่อที่มีครบทุกกลุ่มสาระ อีกทั้งมหาวิทยาลัย ได้ดำเนินการอบรมเพื่อพัฒนาครูและบุคลากรทางการศึกษา ในการสร้างแอปพลิเคชันการเรียนรู้ เพื่อส่งเสริมการพัฒนาครูในการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนการสอนและ

การประยุกต์ใช้สื่ออีดีแอลทีวีเพื่อการเรียนการสอน โดยได้ดำเนินการจัดโครงการอบรมการพัฒนาแอปพลิเคชันสำหรับครูและบุคลากรทางการศึกษา เนื่องจากการจัดการอบรมครั้งนี้มีครูและบุคลากรทางการศึกษาให้ความสนใจสมัครเข้ารับการอบรมเป็นจำนวนมาก และจากการจัดอบรมที่ผ่านมาพบว่า ยังไม่ครอบคลุม เนื่องจากยังมีครูและบุคลากรทางการศึกษาอีกจำนวนมากที่สนใจสมัครเข้ารับการอบรม [3]

จากหลักการและเหตุผลดังกล่าวข้างต้น ผู้วิจัยได้เห็นความสำคัญและความจำเป็นในการพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนแท็บเล็ต กลุ่มสาระการเรียนรู้วิทยาศาสตร์ โดยให้ผู้เข้ารับการอบรมศึกษาเนื้อหาสาระการเรียนรู้จากสื่ออีดีแอลทีวี กลุ่มสาระการเรียนรู้วิทยาศาสตร์ นำสู่การออกแบบแอปพลิเคชัน เพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ต เพื่อให้ครูสามารถสร้างแอปพลิเคชัน ที่ใช้เป็นสื่อในการเรียนการสอนให้กับนักเรียนชั้นประถมศึกษาได้อย่างมีคุณภาพ ส่งผลให้นักเรียนมีสื่อที่สามารถใช้ในการเรียนรู้บนแท็บเล็ตเพิ่มมากขึ้น สอดคล้องกับนโยบายของ สพฐ.

2. วรรณกรรมที่เกี่ยวข้อง

2.1 สื่ออีดีแอลทีวี

สื่ออีดีแอลทีวี หรือ eDLTV (Electronic Distance Learning Television) เป็นสื่อที่พัฒนามาจากการดำเนินงานของโครงการจัดทำเนื้อหาในระบบบลิเอร์นนิ่ง (e-Learning) ของการศึกษาทางไกลผ่านดาวเทียม ร่วมกับโครงการเทคโนโลยีสารสนเทศตามพระราชดำริสมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี โดยศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ หรือ เนคเทค (National Electronics and Computer Technology Center : NECTEC) ได้ดำเนินการโครงการจัดทำเนื้อหาบนระบบบลิเอร์นนิ่ง (e-Learning) ของการศึกษาทางไกลผ่านดาวเทียมโดยการนำเนื้อหาวิดีโอทัศน์การสอนที่ถ่ายทอดจากโรงเรียน วังไกลกังวล อำเภอหัวหิน จังหวัดประจวบคีรีขันธ์ มารวบรวมและจัดเก็บด้วยระบบบริหารจัดการเรียนรู้อีดีแอลสแควร์ (eDL-Square) ซึ่งเป็นระบบที่พัฒนาต่อยอดมาจากระบบเลิร์นสแควร์ (Learn Square) เพื่อให้โรงเรียนในโครงการเทคโนโลยีสารสนเทศตามพระราชดำริฯ สามารถจัดการเรียนการสอนโดยใช้การศึกษาทางไกลผ่านดาวเทียมในรูปแบบสื่ออิเล็กทรอนิกส์เพื่อการเรียนรู้แบบไม่ผ่านเครือข่าย หรือ บลิเอร์

นึ่งแบบออฟไลน์ (Off-Line e-Learning) การเรียนการสอนได้ทั้งแบบเครือข่ายระบบแลน (Local Area Network : LAN) และแบบเครื่องเดี่ยว (Stand Alone) [5]

2.2 การพัฒนาหลักสูตรอบรม

รูปแบบ ADDIE เป็นรูปแบบที่ได้รับการยอมรับกันอย่างกว้างขวางในการนำมาใช้ในการพัฒนาหลักสูตรอบรม โดย ADDIE Model ประกอบด้วย 5 ขั้นตอน ได้แก่ 1. ขั้นตอนวิเคราะห์ (A: Analysis) เป็นขั้นวางแผนหรือเตรียมการสื่อต่างๆ ที่จำเป็นต่อการพัฒนาหลักสูตรอบรม 2. ขั้นตอนออกแบบ (D: Design) เป็นขั้นที่นำข้อมูลต่างๆ ที่ได้วิเคราะห์ไว้เป็นข้อมูลพื้นฐานในการออกแบบ 3. ขั้นตอนพัฒนา (D: Development) เป็นขั้นที่นำสิ่งต่างๆ ที่ได้ออกแบบไว้พัฒนา เมื่อดำเนินการพัฒนาหลักสูตรแล้ว ผู้ออกแบบจะต้องนำหลักสูตรไปทดสอบ เพื่อตรวจสอบหาความผิดพลาด และเพื่อความสมบูรณ์ของแต่ละขั้นตอน 4. ขั้นตอนทดลองใช้ (I: Implementation) เป็นขั้นที่นำหลักสูตรอบรมที่มีความสมบูรณ์มาทดลองใช้ เพื่อหาประสิทธิภาพของหลักสูตรอบรม และ 5. ขั้นตอนประเมินผล (E: Evaluation) โดยการนำผลทดลองที่ได้มาสรุปผล [6]

2.3 งานวิจัยที่เกี่ยวข้อง

อรฉัน พันเดือน [7] ทำวิจัยเรื่อง การพัฒนาหลักสูตรอบรมการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ เรื่อง การทำภาพสติกเกอร์ด้วย iPad2 สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 1 โดยดำเนินการตามขั้นตอน ADDIE Model หลักสูตรอบรมประกอบด้วยเนื้อหา ความรู้พื้นฐานเกี่ยวกับแท็บเล็ต ความรู้พื้นฐานเกี่ยวกับ iPad2 วิธีการใช้งาน iPad2 เบื้องต้น การใช้งานแอปพลิเคชัน บนกล้อง iPad2 และการทำภาพสติกเกอร์ด้วย iPad2 ผลการประเมินหลักสูตรอบรมโดยผู้เชี่ยวชาญมีคุณภาพในระดับเหมาะสมมาก ผู้เข้ารับการอบรมมีความรู้ความเข้าใจหลักสูตรอบรม โดยมีคะแนนหลังอบรมสูงกว่าก่อนอบรมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และมีความพึงพอใจต่อหลักสูตรอบรมอยู่ในระดับมากที่สุด

ประภาวดี กาญจนวิลานนท์ [8] ทำการวิจัยเรื่อง การพัฒนาหลักสูตรอบรมการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ เรื่อง การสร้างสื่อวีดิทัศน์ด้วย iPad2 ชั้นมัธยมศึกษาปีที่ 6 โดยดำเนินการตามขั้นตอน ADDIE Model หลักสูตรอบรมที่พัฒนาขึ้นประกอบด้วยเนื้อหา ความรู้พื้นฐานเกี่ยวกับแท็บ

เล็ด ความรู้พื้นฐานเกี่ยวกับ iPad2 ส่วนประกอบของ iPad2 วิธีใช้งาน iPad2 เบื้องต้น การใช้งานแอปพลิเคชันพื้นฐาน การสร้างสื่อวิดีโอด้วย iMovie ผลการประเมินคุณภาพหลักสูตรอยู่ในระดับเหมาะสมมาก ผู้เข้ารับการอบรมมีความรู้ ความเข้าใจในการอบรมโดยมีคะแนนหลังการอบรมสูงกว่าคะแนนก่อนการอบรมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และผู้เข้ารับการอบรมมีความพึงพอใจต่อหลักสูตรการอบรมที่พัฒนาขึ้นในระดับมาก

นำชัย โบราณมูล [9] ทำการพัฒนาหลักสูตรอบรมการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ เรื่อง การสร้างสเปรดชีตด้วย Numbers สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยดำเนินการตามขั้นตอน ADDIE Model หลักสูตรอบรมประกอบด้วยเนื้อหา ความรู้พื้นฐานการใช้คอมพิวเตอร์แท็บเล็ต ส่วนประกอบของแท็บเล็ต วิธีการใช้แท็บเล็ต การใช้แอปพลิเคชัน การสร้างสเปรดชีตด้วย Numbers และผลการประเมินหลักสูตรมีคุณภาพในระดับมาก ผู้เข้ารับการอบรมมีความรู้ ความเข้าใจในการอบรมโดยมีคะแนนหลังการอบรมสูงกว่าก่อนการอบรมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และผู้เข้ารับการอบรมมีความพึงพอใจต่อรูปแบบการอบรมระดับมาก

Nicholas Gorgievski and Others [10] ได้ทำการศึกษา เรื่อง เครื่องมือสำหรับการเรียนการสอนแคลคูลัส พบว่า การรับรู้ของนักเรียนจากเครื่อง Tablet PC เป็นเครื่องมือในการเรียนการสอนสำหรับการสอนแคลคูลัส ซึ่งวัตถุประสงค์ของการสำรวจครั้งนี้ คือ การเก็บรวบรวมข้อมูลเกี่ยวกับการรับรู้ของนักเรียนเกี่ยวกับการใช้งานของเครื่อง Tablet PC เป็นเครื่องมือในการเรียนการสอนได้อย่างมีประสิทธิภาพและมีประสิทธิภาพครอบคลุมวัสดุการเรียนการสอนในชั้นเรียนผลการชี้ให้เห็นว่านักเรียนรับรู้ว่าเครื่อง Tablet PC มีความสนใจในการเรียนในชั้นเรียน ช่วยให้เข้าใจเนื้อหาในเรื่องที่เรียนครอบคลุมเนื้อหาในการเรียน ทำให้เกิดประสิทธิภาพในการเรียน

Gregory Pryor and Vemon Bauer [11] ได้ทำการทดสอบเทคโนโลยีคอมพิวเตอร์แท็บเล็ตในหลักสูตรการสร้างแล็บชีววิทยา ซึ่งเทคโนโลยีแท็บเล็ตพีซีสามารถทำให้ผู้เรียนสภาพแวดล้อมของห้องเรียนมีชีวิตชีวาเพราะการเรียนด้วยคอมพิวเตอร์แท็บเล็ตเป็นการเรียนแบบพลวัตที่มีการโต้ตอบ

ซึ่งแตกต่างจากเครื่องคอมพิวเตอร์ทั่วไป การใช้คอมพิวเตอร์แท็บเล็ต สามารถทำให้ผู้สอนและนักเรียนใช้เป็นเครื่องมือใช้อธิบาย ชี้แจง และไฮไลต์เอกสารได้โดยตรงจากการใช้งานการสัมผัสจอภาพ ซึ่งไม่เหมือนกับการเรียนจากคอมพิวเตอร์ที่แสดงในชั้นเรียน

จากการศึกษาเอกสารและงานวิจัย พบว่าสื่ออีดีแอลทีวีมีสื่ออิเล็กทรอนิกส์ จำนวนมาก ครูสามารถใช้เป็นเครื่องมือในการเรียนรู้และนำเนื้อหาจากสื่ออีดีแอลทีวีสู่การพัฒนาแอปพลิเคชันบทแท็บเล็ต ให้ผู้เรียนสนใจเรียนรู้และเป็นการเรียนรู้แบบมีชีวิตชีวาได้ตลอดเวลา จากงานวิจัยการพัฒนาหลักสูตรอบรมโดยใช้ ADDIE Model มีขั้นตอนการดำเนินงาน 5 ขั้นตอน ทั้งนี้หลักสูตรอบรมจะเป็นหลักสูตรที่มีคุณภาพ ทำให้ผู้เข้ารับการอบรมมีความรู้และมีความพึงพอใจ จะต้องเป็นหลักสูตรที่ดำเนินการพัฒนาให้สอดคล้องกับความต้องการและบริบทของผู้เข้ารับการอบรม ผู้วิจัยได้นำผลการศึกษาดังกล่าวมาเป็นแนวทางดำเนินการวิจัยในครั้งนี้

3. ขั้นตอนวิธีการดำเนินการวิจัย

การวิจัยได้ดำเนินการตามขั้นตอนของวิธีการเชิงระบบรูปแบบ ADDIE Model ประกอบด้วย 5 ขั้นตอน ได้แก่

3.1 ขั้นตอนการวิเคราะห์ ได้วิเคราะห์เนื้อหา และกิจกรรมในการจัดการอบรมให้สอดคล้องกับบริบทและความต้องการ

3.2 ขั้นตอนการออกแบบ ดำเนินการออกแบบขั้นตอนการอบรม การบูรณาการสื่ออีดีแอลทีวี ผู้การพัฒนาแอปพลิเคชันออกแบบเครื่องมือการวัดและประเมินผล

3.3 ขั้นตอนการพัฒนา ดำเนินการพัฒนาหลักสูตรอบรมและเครื่องมือการวัดและประเมินผลตามขั้นตอนการออกแบบนำไปทดลองใช้เพื่อปรับปรุง และประเมินคุณภาพโดยผู้เชี่ยวชาญ

3.4 ขั้นตอนการทดลองใช้ นำหลักสูตรอบรมพร้อมเครื่องมือที่พัฒนาขึ้นทดลองใช้กับกลุ่มตัวอย่าง ตามขั้นตอนที่ได้ออกแบบไว้

3.5 ขั้นตอนการประเมิน นำผลที่ได้จากการทดลองใช้มาวิเคราะห์และสรุปผลการดำเนินงานตามวัตถุประสงค์การวิจัย

การวิจัยครั้งนี้ ตัวแปรต้นคือ การพัฒนาหลักสูตรการประยุกต์ใช้สื่ออีดีแอลทีวีเพื่อการเรียนรู้สู่การพัฒนาแอปพลิเคชันบนแท็บเล็ต กลุ่มสาระการเรียนรู้วิทยาศาสตร์ที่พัฒนาขึ้น และตัว

แปรตามที่ศึกษา คือ ผลการพัฒนาแอปพลิเคชันของผู้เข้ารับการอบรม และความพึงพอใจของผู้เข้ารับการอบรม

4. ผลการวิจัย

4.1 ผลการพัฒนาหลักสูตรอบรมการประยุกต์ใช้สื่ออีดีแอลทีวีเพื่อการเรียนรู้สู่การพัฒนาแอปพลิเคชันบนแท็บเล็ต กลุ่มสาระการเรียนรู้วิทยาศาสตร์ พบว่า หลักสูตรที่พัฒนาขึ้น มี 6 ชั้น ดังนี้
ชั้นที่ 1 แนะนำสื่ออีดีแอลทีวีและตัวอย่างแอปพลิเคชันเพื่อการเรียนรู้บนแท็บเล็ต โดยใช้สื่ออีดีแอลทีวีและแอปพลิเคชันจากเว็บไซต์ www.otpcappcon.com

ชั้นที่ 2 การเลือกเนื้อหา ผู้เข้ารับการอบรมเลือกสื่อแนะนำเสนอจากสื่ออีดีแอลทีวีกลุ่มสาระการเรียนรู้วิทยาศาสตร์ คนละ 1 เรื่อง ที่สอดคล้องกับบริบทและตามความต้องการ

ชั้นที่ 3 การปรับปรุงเนื้อหา ผู้เข้ารับการอบรมปรับปรุงเนื้อหาจากสื่อแนะนำให้สอดคล้องกับกิจกรรมการเรียนการสอน และนำสู่การพัฒนาเนื้อหาเพื่อการพัฒนาแอปพลิเคชัน ด้วยเครื่องมือ DoToon

ชั้นที่ 4 การจัดทำเกมเพื่อการเรียนรู้ โดยจัดทำเกมจากเนื้อหาที่เลือกประกอบด้วย 5 เกม ได้แก่ เกม Matching เกมแยกประเภท เกมเรียงภาพ เกม Hang man Thai เกมถูก หรือ ผิด

ชั้นที่ 5 การรวบรวมเครื่องมือ เป็นขั้นตอนการนำเนื้อหาและเกมที่สร้างขึ้นมาจัดทำเป็นแอปพลิเคชันและตรวจสอบความถูกต้อง ปรับปรุงแก้ไขและนำขึ้นบนแท็บเล็ต

ชั้นที่ 6 การประเมินแอปพลิเคชันที่พัฒนาขึ้นโดยผู้ทรงคุณวุฒิ จำนวน 3 คน ได้แก่ วิทยากร ศึกษานิเทศก์ และอาจารย์ผู้สอนด้านวิทยาศาสตร์จากมหาวิทยาลัย โดยใช้แบบประเมินที่ผู้วิจัยพัฒนาขึ้น

ผลการประเมินคุณภาพของหลักสูตรอบรมที่พัฒนาขึ้น โดยผู้เชี่ยวชาญจำนวน 5 คน แสดงดังตารางที่ 1

ตารางที่ 1: ผลการประเมินคุณภาพหลักสูตรอบรมโดยผู้เชี่ยวชาญ

รายการประเมิน	ผลการประเมิน		
	\bar{X}	S.D.	ระดับคุณภาพ
1. หลักการและเหตุผล	4.53	0.52	เหมาะสมมากที่สุด
2. วัตถุประสงค์ของหลักสูตรอบรม	4.60	0.50	เหมาะสมมากที่สุด
3. โครงสร้างของหลักสูตรอบรม ทั้ง 6 ชั้น	4.55	0.51	เหมาะสมมากที่สุด
4. ระยะเวลาการอบรม 2 วัน	4.65	0.49	เหมาะสมมากที่สุด

รายการประเมิน	ผลการประเมิน		
	\bar{X}	S.D.	ระดับคุณภาพ
5. การวัดผลการอบรม	4.73	0.46	เหมาะสมมากที่สุด
6. การประเมินผลการอบรม	4.70	0.48	เหมาะสมมากที่สุด
เฉลี่ยรวมทุกด้าน	4.63	0.49	เหมาะสมมากที่สุด

จากตารางที่ 1 พบว่า ผู้เชี่ยวชาญมีความคิดเห็นต่อหลักสูตรอบรมที่พัฒนาขึ้น โดยรวมในระดับเหมาะสมมากที่สุด ($\bar{X} = 4.63, S.D. = 0.49$) เพื่อพิจารณารายด้านพบว่า ทุกด้านมีคุณภาพอยู่ในระดับเหมาะสมมากที่สุด

4.2 ผลการศึกษาความรู้และความเข้าใจของครูที่เข้ารับการอบรมตามหลักสูตรอบรมที่พัฒนาขึ้น ดำเนินการโดยวิเคราะห์เปรียบเทียบคะแนนเฉลี่ยก่อนและหลังการอบรม โดยใช้การทดสอบแบบค่าที่ t-test (Dependent Samples) ดังตารางที่ 2

ตารางที่ 2: ผลการเปรียบเทียบคะแนนการอบรม

กลุ่มตัวอย่าง	คะแนนเต็ม	คะแนนก่อนอบรม		คะแนนหลังอบรม		t
		\bar{X}	S.D.	\bar{X}	S.D.	
		30	40	13.23	9.97	

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตารางที่ 2 พบว่าค่า t ที่คำนวณได้มีค่า 15.41 ซึ่งมีค่ามากกว่าค่า t (ตาราง $\alpha=.05, df=29$) มีค่า 1.699 สรุปได้ว่า ผลการศึกษาความรู้และความเข้าใจของครูที่เข้ารับการอบรมตามหลักสูตรอบรมที่พัฒนาขึ้น ครูที่เข้ารับการอบรมมีคะแนนความรู้และความเข้าใจเฉลี่ยหลังการอบรมสูงกว่าก่อนการอบรม อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

4.3 ผลการศึกษาผลการพัฒนาแอปพลิเคชันของครูที่เข้ารับการอบรมหลักสูตรอบรมที่พัฒนาขึ้น จากผลการประเมินคุณภาพแอปพลิเคชัน โดยผู้เชี่ยวชาญ จำนวน 5 คน พบว่า แอปพลิเคชันของครูที่เข้ารับการอบรมพัฒนาขึ้น มีคุณภาพอยู่ในระดับมากถึงมากที่สุด ($\bar{X} = 4.33- 4.83, S.D. = 0.37- 0.67$)

4.4 ผลการศึกษาความพึงพอใจของครูที่เข้ารับการอบรมหลักสูตรที่พัฒนาขึ้น โดยทำการสอบถามความพึงพอใจของครูที่เข้ารับการอบรมหลังจากจัดการอบรมตามขั้นตอนทั้ง 6 ชั้น ผลการประเมิน แสดงดังตารางที่ 3

ตารางที่ 3: ความพึงพอใจของครูที่เข้ารับการอบรม

รายการ	ความพึงพอใจ		
	\bar{X}	S.D.	ระดับความพึงพอใจ
1. ด้านวิทยากร	4.72	0.45	มากที่สุด
2. ด้านกระบวนการอบรม	4.56	0.50	มากที่สุด
3. ด้านความรู้ความเข้าใจ	4.62	0.49	มากที่สุด
4. ด้านการนำไปใช้ประโยชน์	4.82	0.39	มากที่สุด
5. ด้านสถานที่และระยะเวลา	4.73	0.48	มากที่สุด
เฉลี่ยรวมทุกด้าน	4.69	0.47	มากที่สุด

จากตารางที่ 3 พบว่า ความพึงพอใจของครูที่เข้ารับการอบรมตามหลักสูตรที่พัฒนาขึ้น เฉลี่ยรวมทุกด้านมีความพึงพอใจอยู่ในระดับมากที่สุด ($\bar{X} = 4.69, S.D. = 0.47$)

5. สรุปผลการวิจัย

การพัฒนาหลักสูตรอบรมการประยุกต์ใช้สื่ออีดีแอลทีวีเพื่อการเรียนรู้สู่การพัฒนาแอปพลิเคชันบนแท็บเล็ต กลุ่มสาระการเรียนรู้วิทยาศาสตร์ หลักสูตรอบรมที่ผู้วิจัยพัฒนาขึ้นประกอบด้วย 6 ชั้น คือ ชั้นที่ 1 แนะนำสื่ออีดีแอลทีวีและตัวอย่างแอปพลิเคชันเพื่อการเรียนรู้บนแท็บเล็ต ชั้นที่ 2 การเลือกเนื้อหา ชั้นที่ 3 การปรับปรุงเนื้อหา ชั้นที่ 4 การจัดทำเกมเพื่อการเรียนรู้ ชั้นที่ 5 การรวบรวมเครื่องมือ และ ชั้นที่ 6 การประเมินแอปพลิเคชันที่พัฒนาขึ้น และผลการประเมินคุณภาพโดยผู้เชี่ยวชาญอยู่ในระดับเหมาะสมมากที่สุด ผู้เข้าอบรมตามหลักสูตรที่พัฒนาขึ้นมีความรู้และเข้าใจโดยมีคะแนนหลังการอบรมสูงกว่าก่อนการอบรมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ผลการศึกษากการพัฒนาแอปพลิเคชันที่ผู้เข้ารับการอบรมพัฒนาขึ้นมีคุณภาพอยู่ในระดับความเหมาะสมมาก ถึง มากที่สุด และผู้เข้ารับการอบรมมีความพึงพอใจต่อหลักสูตรอบรมอยู่ในระดับมากที่สุด

6. กิตติกรรมประกาศ

การวิจัยครั้งนี้ดำเนินการบรรลุตตามวัตถุประสงค์ ผู้วิจัยขอขอบคุณ สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ หรือ สวทช. โดยศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ หรือเนคเทค (NECTEC) มูลนิธิการศึกษาทางไกลผ่านดาวเทียม โครงการเทคโนโลยีสารสนเทศตาม

พระราชดำริสมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี และ วิทยากรจากคณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏมหาสารคาม และครูผู้สอนกลุ่มสาระการเรียนรู้วิทยาศาสตร์ ชั้นประถมศึกษาปีที่ 1 สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษามหาสารคาม เขต 1

7. เอกสารอ้างอิง

- [1] กระทรวงศึกษาธิการ. โครงการแท็บเล็ตพีซีเพื่อการศึกษาไทย. [ออนไลน์].สืบค้นจาก <http://www.otpc.in.th/aboutus.html> (เมื่อ 17 มีนาคม 2556)
- [2] ราชภัฏมหาสารคาม, มหาวิทยาลัย. รายงานผลการดำเนินงานการส่งเสริมการพัฒนาครู บุคลากรทางการศึกษา และนักเรียนในการใช้เทคโนโลยีสารสนเทศเพื่อยกระดับการศึกษาเรียนรู้ของชุมชน มหาวิทยาลัยราชภัฏมหาสารคาม. มหาสารคาม : มหาวิทยาลัยราชภัฏมหาสารคาม, 2555.
- [3] สมเจตน์ ภูศรี “รายงานการดำเนินงานเครือข่ายเผยแพร่ถ่ายทอดและพัฒนาสื่อการเรียนการสอนบนระบบ e-Learning (eDL- Square)”. มหาวิทยาลัยราชภัฏมหาสารคาม, 2555.
- [4] กระทรวงศึกษาธิการ. ตัวชี้วัดและสาระการเรียนรู้แกนกลาง กลุ่มสาระการเรียนรู้วิทยาศาสตร์ ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐานพุทธศักราช 2551. กรุงเทพฯ: โรงพิมพ์สหกรณ์การเกษตร แห่งประเทศไทย จำกัด, 2551.
- [5] เครือข่ายเผยแพร่และถ่ายทอดการใช้สื่อ eDLTV ในการเรียนการสอน มหาวิทยาลัยราชภัฏมหาสารคาม. [ออนไลน์].สืบค้นจาก <http://www.itrmu.net/edltv/> (เมื่อ 20 มีนาคม 2556)
- [6] พิสุทธิ อาธิราษฎร์. การพัฒนาซอฟต์แวร์ทางการศึกษา. มหาสารคาม : มหาวิทยาลัยราชภัฏมหาสารคาม, 2551.
- [7] อรฉัตร เตือนพั้ง. การพัฒนาหลักสูตรอบรมการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ เรื่อง การทำภาพสติกเกอร์ ด้วย iPad 2. มหาวิทยาลัยราชภัฏมหาสารคาม, 2555.
- [8] ประภาวดี กาญจนวิลานนท์. การพัฒนาหลักสูตรอบรมการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ เรื่อง การสร้างสื่อวีดิทัศน์ด้วย iPad2 สำหรับชั้นมัธยมศึกษาปีที่ 6. มหาวิทยาลัยราชภัฏมหาสารคาม, 2555.
- [9] นำชัย โบราณมูล. การพัฒนาหลักสูตรอบรมการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ เรื่องการสร้างสเปรดชีตด้วย Numbers สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3. มหาวิทยาลัยราชภัฏมหาสารคาม, 2555.

- [10] Nicholas Gorgievski and other. "Tablet PC: A Preliminary Report on a Tool for Teaching Calculus". *The International Journal for Technology in Mathematics Education*, 95-102, 2005.
- [11] Gregory Pryor and Vemon Bauer. "Buil a Better Biology Lab Testing Tablet PC Techonology in a Core Laboratory Course". *Journal of College Science Teaching*, November/December, 44-48, 2008.

การศึกษากรอบการดำเนินงานกรีนไอทีสำหรับมหาวิทยาลัยราชภัฏ

A Study of Green IT Framework for Rajabhat University

ธรัช อารีราษฎร์¹ ละอองทิพย์ มัทธูรส²

มนต์ชัย เทียนทอง³ และ ดุสนี สุภวรรธนะกุล²

¹นักศึกษาลัทธิศึกษาศาสตร์บัณฑิต สาขาการจัดการเทคโนโลยี มหาวิทยาลัยราชภัฏพระนคร

²คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏพระนคร

³คณะครุศาสตร์อุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

dr.tharach@hotmail.com, laongtip77@hotmail.com, drmonchai@gmail.com, s.dusanee@hotmail.com

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อศึกษากรอบการดำเนินงานกรีนไอทีสำหรับมหาวิทยาลัยราชภัฏ กลุ่มเป้าหมาย ได้แก่ ผู้เชี่ยวชาญจากมหาวิทยาลัยต่าง ๆ จำนวน 5 คน เครื่องมือที่ใช้ ได้แก่ แบบสอบถามความคิดเห็นความเหมาะสมของกรอบการดำเนินงานกรีนไอทีสำหรับมหาวิทยาลัยราชภัฏ สถิติที่ใช้ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน

ผลการวิจัย พบว่า กรอบการดำเนินงานกรีนไอทีสำหรับมหาวิทยาลัยราชภัฏ มี 5 องค์ประกอบ ได้แก่ 1) ด้านนโยบาย 2) ด้านทัศนคติ 3) ด้านเทคโนโลยี 4) การปฏิบัติการ และ 5) การประเมินผล

การดำเนินงานกรีนไอทีที่จัดแบ่งเป็น 3 ระดับ ได้แก่ 1) ระดับส่วนบุคคล 2) ระดับหน่วยงาน และ 3) ระดับมหาวิทยาลัย ผลการศึกษาความคิดเห็นพบว่า ผู้เชี่ยวชาญมีความเห็นโดยรวมในระดับมาก ($\bar{X}=4.40$ และ $SD. = 0.50$) เมื่อพิจารณาองค์ประกอบพบว่า ผู้เชี่ยวชาญมีความเห็นในระดับมากที่สุด ($\bar{X}=4.52$ และ $SD. = 0.51$) และด้านระดับในการดำเนินงานพบว่า ผู้เชี่ยวชาญมีความเห็นในระดับมาก ($\bar{X}=4.20$ และ $SD. = 0.41$)

คำสำคัญ: กรีนไอที ดาตาเซ็นเตอร์ คลาวด์คอมพิวติ้ง ระบบสารสนเทศเพื่อการบริหาร ระบบเสมือน

Abstract

The objective of the research was to study the Green IT Framework for Rajabhat University by using the analysis of related research papers. The target population were five experts in information technology and related. The instrument was an assessment form of Green IT Framework for Rajabhat University. The statistics used was mean and standard deviation.

The results showed that the Green IT framework consists of five components, 1) Policy 2) Attitude 3) Technology 4) Practice and 5) Evaluation.

The operations could be classified into three levels. 1) Personal Level 2) Faculty/Department/Center Level and 3) University Level.

Regarding the results assessed by five experts, the finding indicated that the average level of opinion of experts towards the Green It Framework was high ($\bar{X}=4.40$ and $SD. = 0.50$). The average level of opinion of experts towards the Green It components was very high ($\bar{X}=4.52$ and $SD. = 0.51$). The average level of opinion of experts towards the classified operations was high ($\bar{X}=4.20$ and $SD. = 0.41$)

Keyword: Green IT, Data Center, Cloud Computing, MIS, Virtualization.

1. บทนำ

สองทศวรรษที่ผ่านมาเทคโนโลยีสารสนเทศได้เข้ามามีบทบาทในสังคมไทยส่งผลให้เกิดความเปลี่ยนแปลงในการดำรงชีวิตทำให้เกิดรูปแบบและโอกาสทางการศึกษาดัดต้นทุนในกระบวนการผลิตทางด้านอุตสาหกรรมและเพิ่มศักยภาพในการแข่งขันทางธุรกิจและบริการทั้งภาครัฐและเอกชนการพัฒนาาระบบสารสนเทศเพื่อสนับสนุนการดำเนินงานกิจกรรมดังกล่าวให้มีประสิทธิภาพนั้นนักพัฒนาได้ให้ความสำคัญทางด้านประโยชน์การใช้งานเสถียรภาพของระบบความปลอดภัยของข้อมูลและความสามารถเพิ่มขยายของระบบเท่านั้นแต่ไม่ได้คำนึงถึงปัจจัยทางด้านสิ่งแวดล้อมจาก

การศึกษาข้อมูล พบว่า เทคโนโลยีสารสนเทศนอกจากมีประโยชน์แล้วแต่ถ้าการประยุกต์ใช้งานภายใต้การจัดการที่ไม่มีประสิทธิภาพที่ดีพอเทคโนโลยีสารสนเทศก็จะเป็นต้นเหตุของการเกิดผลกระทบต่อสิ่งแวดล้อมได้ [1] ผลการวิจัยของหน่วยงานบริหารงานข้อมูลทางด้านพลังงานหรือ EPA แสดงให้เห็นว่าการปล่อยก๊าซเรือนกระจก (Greenhouse gases) ร้อยละ 87 ของการปล่อยก๊าซนั้นเป็นคาร์บอนไดออกไซด์ (CO₂) ซึ่งกิจกรรมที่เกิดจากการใช้พลังงานในชีวิตประจำวันของคนเรานั้นส่งผลให้เกิดการปล่อยก๊าซคาร์บอนไดออกไซด์ถึงร้อยละ 98 โดยที่ร้อยละ 2 เกิดจากกิจกรรมทางด้านเทคโนโลยีสารสนเทศประกอบไปด้วยการใช้งานพีซีและจอภาพร้อยละ 39 เซิร์ฟเวอร์ร้อยละ 23 การสื่อสารแบบสายร้อยละ 15 การสื่อสารแบบไร้สายร้อยละ 9 ระบบเครือข่ายในสำนักงานร้อยละ 7 และเครื่องพิมพ์ร้อยละ 6 [1] นอกจากนี้ยังมีขยะอิเล็กทรอนิกส์ (e-Waste) เกิดขึ้น เมื่อขยะเหล่านี้ไม่มีการจัดการที่เหมาะสม จะปลดปล่อยสารปรอทในรูปแบบไอระเหยออกมา และอาจมีสารตะกั่ว สารแคดเมียมด้วย [2] ซึ่งสารเหล่านี้ล้วนเป็นพิษต่อร่างกายของมนุษย์

จากสภาพการณ์ดังกล่าว แนวทางการพัฒนาเทคโนโลยีสารสนเทศให้เป็นมิตรกับสิ่งแวดล้อมจึงได้รับความสนใจศึกษาจากทั่วโลกภายใต้คำนิยามใหม่ที่เรียกว่าเทคโนโลยีสารสนเทศสีเขียว หรือ กรีนไอที (Green IT) [1] ตัวอย่างหน่วยงาน ทั้งภาครัฐและเอกชน ที่ได้ดำเนินการกรีนไอที เพื่อรักษาสีสิ่งแวดล้อม โดยบริษัทผู้ผลิตสินค้าไอทีได้แก่ บริษัท เฮชพี (Hewlett Packard) ผลิตดัลล์มิกเลเซอร์สี และชาวคำ (ชนิดเทียบ และริเมนู) ที่ลดค่าใช้จ่ายลดปริมาณขยะลดโลกร้อนบริษัท ฟุจิซี (Fujitsu) เสนอคอมพิวเตอร์กินไฟต่ำ รุ่นบาง เรียกว่า Fujitsu Futro S100 Thin Client เป็นต้น ในส่วนของสถาบันการศึกษา ได้มีการนำเครื่องคอมพิวเตอร์และอุปกรณ์ด้านเทคโนโลยีสารสนเทศอื่น ๆ มาใช้เป็นเครื่องมือในงานบริหาร และการจัดการเรียนการสอน กันอย่างกว้างขวาง และมีแนวโน้มจะเพิ่มมากขึ้น

สำหรับการดำเนินงานด้านกรีนไอที ในมหาวิทยาลัยหลาย ๆ แห่ง ได้ดำเนินการแล้ว เช่น มหาวิทยาลัยมิชิแกน [3] มหาวิทยาลัยคาร์ดิฟ (Cardiff University) [4] หรือมหาวิทยาลัยในประเทศไทยเกือบทุกแห่งได้บรรจุเป็นมาตรการหรือกลยุทธ์ในแผนยุทธศาสตร์ของมหาวิทยาลัยเช่น จุฬาลงกรณ์มหาวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์ มหาวิทยาลัยแม่โจ้ หรือมหาวิทยาลัยราชภัฏมหาสารคาม เป็นต้น

บุคลากรในมหาวิทยาลัยราชภัฏ ประกอบไปด้วยบุคลากรสายวิชาการ สายสนับสนุน ตลอดจนนักศึกษา ซึ่งมีความรู้ในด้าน กรีนไอทีที่ไม่เท่ากัน ส่วนใหญ่ไม่รู้เกี่ยวกับกรีน

ไอที เนื่องจากไม่ได้สำเร็จการศึกษาด้านเทคโนโลยีสารสนเทศ หรือภาระหน้าที่รับผิดชอบมีมากจน ไม่มีเวลาที่จะศึกษาหาความรู้เพิ่มเติมด้านไอที หรือหน่วยงานไม่มีมาตรการด้านไอที จึงไม่มีความสนใจและไม่มีความปฏิบัติด้านกรีนไอที นอกจากนี้วัฒนธรรมองค์กร หรือบุคคล ในการใช้งานไอที จะพบว่า บุคลากร มีการเปิดเครื่องคอมพิวเตอร์ไว้โดยไม่ใช้งาน ในช่วงเวลากลางวัน และมีการเปิดเครื่องไว้โดยไม่ใช้งานในช่วงเวลากลางคืน ในส่วนของนักศึกษาพบว่า ไม่ได้ปิดเครื่องหลังจากเข้าไปใช้ในห้องปฏิบัติการคอมพิวเตอร์

จากความจำเป็นในการลดการใช้พลังงาน จากการใช้งานด้านคอมพิวเตอร์ของบุคลากร และนักศึกษา ผู้วิจัยจึงสนใจศึกษารอบการดำเนินงานกรีนไอทีสำหรับมหาวิทยาลัยราชภัฏ เพื่อเป็นแนวทางในการดำเนินการกรีนไอทีสำหรับมหาวิทยาลัยราชภัฏต่อไป

วัตถุประสงค์

เพื่อศึกษารอบการดำเนินงานกรีนไอทีสำหรับมหาวิทยาลัยราชภัฏ

2. วรรณกรรมที่เกี่ยวข้อง

Murugesan และ Harnessing [9] กำหนดกรอบการดำเนินงานกรีนไอทีใน 4 ด้าน โดยประกอบด้วย 4 องค์ประกอบ ได้แก่ 1) Green Disposal เป็นหลักการในการนำคอมพิวเตอร์เก่ามาใช้งานใหม่และทำการรีไซเคิล (Recycle) คอมพิวเตอร์และอุปกรณ์อิเล็กทรอนิกส์ที่ไม่ต้องการอย่างเหมาะสม 2) Green Design เป็นหลักการในการออกแบบอุปกรณ์คอมพิวเตอร์ คอมพิวเตอร์แม่ข่ายและอุปกรณ์ทำความเย็นโดยประหยัดพลังงานและรักษาสีสิ่งแวดล้อม 3) Green Manufacturing เป็นหลักการในการประดิษฐ์และสร้างอุปกรณ์อิเล็กทรอนิกส์คอมพิวเตอร์และระบบย่อยอื่น ๆ ที่มีส่วนเชื่อมโยงกันโดยมีผลกระทบต่อสิ่งแวดล้อมน้อยที่สุดหรือไม่มีเลย และ 4) Green Use เป็นหลักการในการใช้ไอทีโดยมีเป้าหมายเพื่อลดการใช้พลังงาน

Open Green IT Policy [8] ได้กำหนดกรอบของกรีนไอที โดยมีองค์ประกอบจำนวน 4 ด้าน ได้แก่ 1) ด้านทัศนคติ/วัฒนธรรมองค์กร (Attitude/Culture) 2) ด้านนโยบายขององค์กร (Policy) 3) ด้านแนวปฏิบัติ (Practice) และ 4) ด้านเทคโนโลยี (Technology) โดยที่แต่ละองค์ประกอบจะประกอบด้วยหน่วยย่อยที่เหมือนกัน ได้แก่ 1) กระบวนการและการกำกับดูแล (Processes & Governance) 2) วงจรชีวิตของอุปกรณ์ (Procurement & Equipment lifecycle) 3) สภาพ

แวดล้อมของสำนักงาน (Office Environment : end user computing) เช่น การใช้คอมพิวเตอร์ของบุคลากร เป็นต้น 4) การใช้ศูนย์ข้อมูล (Data Center) 5) ตัวชี้วัดและการรายงาน (Metrics & Reporting) และ 6) การใช้ไอทีช่วยลดปริมาณคาร์บอน (IT as low Carbon enabler)

Graem Philipson [5] ได้กำหนดกรอบของการดำเนินงานกรีนไอที โดยมีองค์ประกอบจำนวน 5 ด้าน ได้แก่ 1) ด้านทัศนคติ (Attitude) 2) ด้านนโยบายขององค์กร (Policy) 3) ด้านแนวปฏิบัติ (Practice) 4) ด้านเทคโนโลยี (Technology) และ 5) ตัวชี้วัด (Metrics) โดยที่แต่ละองค์ประกอบจะประกอบด้วยหน่วยย่อยที่เหมือนกัน ได้แก่ 1) ประสิทธิภาพการใช้งานไอทีของบุคลากร (End User IT Efficiencies) ทั้งแบบตั้งโต๊ะ (Desktop) แบบเคลื่อนที่ (Mobile Devices) และการพิมพ์งาน (Printing & Consumables) 2) ประสิทธิภาพการใช้งานไอทีขององค์กร (Enterprise IT Efficiencies) ประกอบด้วย อุปกรณ์ของศูนย์ข้อมูล (Data Center IT Equipment) สภาพแวดล้อมและการใช้พลังงานของศูนย์ข้อมูล (Data Center Power & Environment) การสื่อสารข้อมูล (Communication) การกระจายระบบไอที (Distributed & Dept IT Systems) และการใช้ระบบผู้รับงานจากนอกหน่วยงาน (Outsourced IT Systems) 3) วงจรอายุการใช้งานและการจัดหา (Lifecycle & Procurement) ประกอบด้วย การจัดหา (Procurement) การดำเนินงาน (Operation) และการจัดการอุปกรณ์ (Disposal) และ 4) การใช้ไอทีช่วยลดคาร์บอน (IT as Low-Carbon Enabler) ประกอบด้วย การปฏิบัติและการกำกับดูแล (Compliance & Governance) การทำงานแบบทางไกล (Teleworking) การจัดการกระบวนการทางธุรกิจ (Business Project Management) การจัดการห่วงโซ่อุปทาน (Supply Chain Management) และ ประสิทธิภาพของอุตสาหกรรมเฉพาะ (Industry-specific Efficiencies) อย่างไรก็ตาม ได้มีการปรับกรอบของการดำเนินงานกรีนไอทีใหม่ [6] โดยปรับเป็น 2 หน่วยย่อย ได้แก่ 1) ด้านการใช้งานไอทีของบุคลากร ปรับใหม่ เป็น การใช้งานระดับบุคคล (Personal Computing) ระบบแผนก (Department Computing) และ 2) ด้านวัฏจักรชีวิตของอุปกรณ์ ปรับใหม่เป็นการนำกลับไปใช้ใหม่ (Recycle & Reuse)

Green IT Foundation [7] ได้กำหนดกรอบของการดำเนินงานกรีนไอที ดังแสดงในภาพที่ 5 โดยมีองค์ประกอบจำนวน 4 ด้าน ได้แก่ 1) ด้านบุคลากร (People) 2) ด้านกระบวนการ (Process) 3) ด้านเทคโนโลยี (Technology) และ 4) ด้านมาตรการ การวัด และการจัดการ (Measure & Manage) ทั้งนี้

แต่ละองค์ประกอบ ประกอบด้วยส่วนย่อยที่เกี่ยวกับผู้ใช้งาน (Consumer/User Optimization) ซึ่งมีรายละเอียดในด้านการใช้คอมพิวเตอร์ในระดับองค์กร ระดับบุคคล และเกี่ยวกับวัสดุ ความสิ้นเปลืองต่าง ๆ สำหรับส่วนย่อยต่อมา ได้แก่ โครงสร้างพื้นฐาน (Infrastructure service Optimization) ได้แก่ ศูนย์ข้อมูล เครือข่ายการสื่อสาร สภาพแวดล้อมโครงสร้างพื้นฐานด้านการจัดการ เช่น การใช้หน่วยงานภายนอก (Outsource) การใช้คลาวด์ (Cloud) เป็นต้น

3. ขั้นตอนวิธีการดำเนินการวิจัย

3.1 ขั้นตอนวิธีการดำเนินการวิจัย การวิจัยครั้งนี้เป็นการวิจัยเชิงคุณภาพและปริมาณ โดยการวิเคราะห์จากเอกสารงานวิจัยที่เกี่ยวข้อง มีขั้นตอนการวิจัยดังนี้

3.1.1 ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องดังนี้

- 1) หลักการด้านกรีนไอที
- 2) หลักการด้านเทคโนโลยีสารสนเทศ ในการดำเนินงานกรีนไอที
- 3) เอกสารงานวิจัยด้านกรอบการดำเนินงานกรีนไอที
- 4) แผนเทคโนโลยีสารสนเทศของมหาวิทยาลัยราชภัฏต่าง ๆ

3.1.2 สรุปผลจากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง

3.1.3 สังเคราะห์กรอบการดำเนินงานกรีนไอทีสำหรับมหาวิทยาลัยราชภัฏ โดยดำเนินการดังนี้

- 1) จัดกลุ่มองค์ประกอบของกรอบการดำเนินงานกรีนไอทีจากเอกสารงานวิจัยที่เกี่ยวข้อง [5] [6] [7] [8] [9]
- 2) จัดทำตาราง (Square matrix) เพื่อสังเคราะห์องค์ประกอบของกรอบการดำเนินงานกรีนไอที
- 3) ร่างกรอบการดำเนินงานกรีนไอทีสำหรับมหาวิทยาลัยราชภัฏ
- 4) นำเสนออาจารย์ที่ปรึกษา เพื่อปรับปรุงแก้ไข

3.1.4 นำเสนอผู้เชี่ยวชาญ จำนวน 5 คน เพื่อศึกษาความคิดเห็นในด้านความเหมาะสมของ กรอบการดำเนินงานกรีนไอทีสำหรับมหาวิทยาลัยราชภัฏ

3.2 กลุ่มเป้าหมาย ในการวิจัยครั้งนี้ ได้แก่ ผู้เชี่ยวชาญจากมหาวิทยาลัยต่าง ๆ จำนวน 5 คน ที่สำเร็จการศึกษาระดับปริญญาเอก สาขาเทคโนโลยีสารสนเทศ หรือที่สัมพันธ์

3.3 เครื่องมือที่ใช้ แบบสอบถามความคิดเห็นความเหมาะสมของกรอบการดำเนินงานกรีนไอทีสำหรับมหาวิทยาลัยราชภัฏ

3.4 สถิติที่ใช้ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน

4. ผลการดำเนินการวิจัย

4.1 ผลการศึกษาเอกสารงานวิจัยที่เกี่ยวข้อง ผู้วิจัยได้

สรุป ดังแสดงในตารางที่ 1

ตารางที่ 1 : ผลการศึกษาเอกสาร

ประเด็นศึกษา	ผลสรุป
หลักการด้านกรีน ไอที [9]	กรอบการดำเนินงานกรีน ไอที ประกอบด้วย 4 องค์ประกอบ ได้แก่ 1) Green Disposal 2) Green Design 3) Green Use และ 4) Green Manufacturing
หลักการด้านเทคโนโลยีสารสนเทศ ในการดำเนินงานกรีน ไอที [1][2]	ประกอบด้วย 1)เทคโนโลยีแบบเสมือน 2)ศูนย์ข้อมูล 3) ระบบคลาวด์คอมพิวเตอร์ 4) ระบบสารสนเทศเพื่อการบริหาร 5)การทำงานและการประชุมทางไกล 6) อุปกรณ์ประหยัดพลังงาน
งานวิจัยด้านกรอบการดำเนินงานกรีน ไอที [5] [6] [7] [8] [9]	ประกอบด้วย 5 องค์ประกอบที่สำคัญ ได้แก่ Policy Attitude Technology Practice Metrics
แผนเทคโนโลยีสารสนเทศของมหาวิทยาลัยราชภัฏต่าง ๆ	มหาวิทยาลัยราชภัฏ จัดทำแผนเทคโนโลยีสารสนเทศในระยะ 5 ปี โดยมียุทธศาสตร์ และกลยุทธ์สำหรับการดำเนินงานกรีน ไอที

4.2 ผลการจัดกลุ่มองค์ประกอบและการสังเคราะห์องค์ประกอบของกรอบการดำเนินงานกรีน ไอที ผลการดำเนินงาน มีรายละเอียดดังนี้

4.2.1 ผู้วิจัยจัดกลุ่มองค์ประกอบจากเอกสารงานวิจัยที่เกี่ยวข้อง [5] [6] [7] [8] [9] ได้จำนวน 5 กลุ่ม ดังนี้

- 1) People, attitude, culture เป็นกลุ่มที่เกี่ยวกับบุคลากร ทักษะคิดและวัฒนธรรมของบุคลากร ในด้านกรีน ไอที
- 2) Process, practice เป็นกลุ่มที่เกี่ยวข้องกับกระบวนการหรือแนวทางการปฏิบัติกรีน ไอที
- 3) Technology เป็นเทคโนโลยีที่ใช้ในการดำเนินงานด้านกรีน ไอที
- 4) Measure and manage, metrics เป็น

กลุ่มที่เกี่ยวกับการจัดการวัดประเมินผลการดำเนินงานกรีน ไอที ซึ่งอาจจำเป็นต้องมีตัวชี้วัดสำหรับการประเมิน

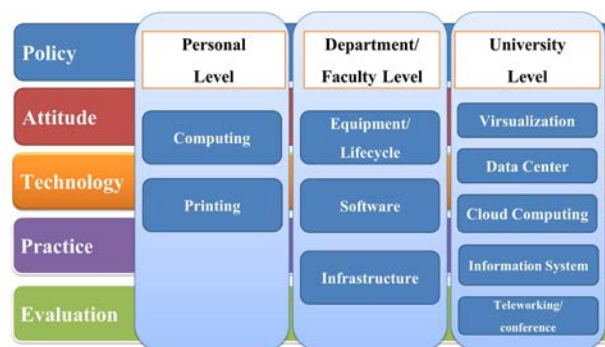
5) Policy เป็นนโยบายของมหาวิทยาลัยที่ใช้เป็นแนวทางในดำเนินงานกรีน ไอที

4.2.2 ผลการสังเคราะห์องค์ประกอบของกรอบการดำเนินงานกรีน ไอที โดยผู้วิจัยนำองค์ประกอบที่จัดกลุ่มตามหัวข้อ 4.2.1 มาเทียบกับผลการศึกษาของนักวิชาการและสถาบันต่าง ๆ [6] [7] [8] [9] โดยการจัดทำตาราง (Square matrix) ผลที่ได้ดังตารางที่ 2

ตารางที่ 2 : การสังเคราะห์องค์ประกอบ

Component	Institution/ Academician		
	Open Green IT Policy	Green Philipson	Green IT Foundation
People, attitude, culture	✓	✓	✓
Process, practice	✓	✓	✓
Technology	✓	✓	✓
Measure and manage, metrics	✓	✓	✓
Policy		✓	✓

4.3 ผลการสังเคราะห์กรอบการดำเนินงานกรีน ไอทีสำหรับมหาวิทยาลัยราชภัฏ ผู้วิจัยนำกลุ่มองค์ประกอบจากตารางที่ 2 เขียนแผนภาพกรอบการดำเนินงานกรีน ไอที รายละเอียดดังแสดงในภาพที่ 1



ภาพที่ 1 : กรอบการดำเนินงานกรีน ไอทีสำหรับมหาวิทยาลัยราชภัฏ

จากภาพที่ 1 กรอบการดำเนินงานกรีน ไอทีสำหรับมหาวิทยาลัยราชภัฏ มี 2 ส่วนคือ

4.3.1 ด้านองค์ประกอบ มี 5 องค์ประกอบดังนี้

- 1) ด้านนโยบาย (Policy) มหาวิทยาลัยราชภัฏจะต้องมีนโยบายที่ชัดเจนในด้านการดำเนินงานกรีน ไอที เพื่อ

เป็นกรอบในการดำเนินงานของหน่วยงานที่เกี่ยวข้อง และบุคลากรทั้งมหาวิทยาลัย ถ่ายทอดนโยบายลงสู่หน่วยงานและบุคลากร ให้ความสำคัญโดยทั่วกันเพื่อถือปฏิบัติ ในขณะที่เดียวกัน มหาวิทยาลัย ต้องจัดสรรงบประมาณให้กับหน่วยงานที่เกี่ยวข้องสำหรับการดำเนินงานกรีนไอที

2) ด้านทัศนคติ (Attitude) บุคลากรจะต้องมีทัศนคติที่ดีหรือมีจิตสำนึกที่ดีต่อกรีนไอที มีความรู้ ความเข้าใจ ถึงของการใช้งานเทคโนโลยีสารสนเทศที่ส่งผลกระทบต่อการใช้พลังงาน และสิ่งแวดล้อม ถ้าบุคลากรยังขาดความรู้ ความเข้าใจ จึงจำเป็นต้องสร้างความตระหนัก ความรู้ ความเข้าใจ ถึงผลกระทบต่าง ๆ ให้กับบุคลากร

3) ด้านเทคโนโลยี (Technology) มหาวิทยาลัยราชภัฏ จะต้องมีการปรับเปลี่ยนเทคโนโลยีแบบค่อยเป็นค่อยไป หรือ ปรับใหม่ทั้งหมดในทันที มีรายละเอียดหรือคุณลักษณะด้านเทคโนโลยีที่เหมาะสมในการดำเนินงานกรีนไอที ที่พร้อมจะถ่ายทอดลงสู่หน่วยงานหรือบุคลากร ตลอดจนมีเทคโนโลยีที่พร้อมสำหรับการดำเนินงานกรีนไอที เพื่อให้หน่วยงานในระดับคณะ/สำนัก/ศูนย์/สาขาวิชา และระดับบุคคล สามารถนำไปปรับใช้หรือบริหารจัดการ

4) การปฏิบัติการ (Practice) มหาวิทยาลัยราชภัฏ จะต้องมีการปฏิบัติ การ มาตรการ เทคโนโลยีที่ใช้ ถ่ายทอดลงสู่หน่วยงานหรือบุคลากร เพื่อให้เกิดการปฏิบัติจริง ทั้งในระดับส่วนบุคคล ระดับหน่วยงาน และระดับมหาวิทยาลัย

5) การประเมินผล (Evaluation) มหาวิทยาลัยราชภัฏ จะต้องกำกับ ดูแล การดำเนินงานกรีนไอที มีการประเมินผลการดำเนินงานกรีนไอที โดยอาจมีตัวชี้วัดต่าง ๆ ที่ครอบคลุมทุกองค์ประกอบ และทุกระดับ เป็นเกณฑ์ในการประเมินผล จัดทำรายงานเสนอต่อฝ่ายบริหาร หน่วยงานและบุคลากรในมหาวิทยาลัยรับทราบโดยทั่วกัน

4.3.2 ระดับในการดำเนินงาน จัดแบ่งเป็น 3 ระดับ ดังนี้

1) ระดับส่วนบุคคล (Personal Level)

การใช้งานเครื่องคอมพิวเตอร์ส่วนบุคคลของบุคลากร จะเกี่ยวข้องทั้งในแบบเครื่องตั้งโต๊ะ (Desktop) และในแบบเคลื่อนที่ (Mobile) ทั้งที่เป็นโน้ตบุ๊ก แท็บเล็ต หรือ สมาร์ทโฟน ซึ่งในการดำเนินงาน มหาวิทยาลัยราชภัฏจะต้องมีแนวปฏิบัติที่ดีในการใช้ตามแผนการดำเนินงานกรีนไอที โดยจะต้องมีการให้ความรู้ ความเข้าใจและสร้างความตระหนักหรือสร้างทัศนคติที่ดีก่อนเริ่มดำเนินการ เพื่อให้เกิดประสิทธิภาพให้มากที่สุด

ในส่วนของการพิมพ์เอกสาร (Printing) มหาวิทยาลัยราชภัฏจะต้องมีแนวปฏิบัติที่ดีในการใช้เครื่องพิมพ์ ตามแผนการดำเนินงานกรีนไอที เช่น การลดการใช้กระดาษ การใช้เครื่องพิมพ์ร่วมกัน การใช้หมึกที่ช่วยลดปริมาณคาร์บอน เป็นต้น

2) ระดับหน่วยงาน (Department/Faculty Level) ได้แก่ คณะ/สำนัก/ศูนย์/สาขาวิชา ประกอบด้วยการดำเนินการ ดังนี้

2.1) อุปกรณ์และอายุการใช้งาน (Equipment/Lifecycle) เกี่ยวข้องในด้านต่อไปนี้

2.1.2) แนวปฏิบัติในการใช้อุปกรณ์ เช่น การมีแนวปฏิบัติในการใช้เครื่องพิมพ์ร่วมกันของบุคลากรในหน่วยงาน เป็นต้น

2.1.3) แนวปฏิบัติในการจัดหาอุปกรณ์ เช่น การจัดหาอุปกรณ์ที่เป็นแบบ Multi Functional หรือ อุปกรณ์ที่ช่วยลดการใช้พลังงาน เป็นต้น

2.1.4) แนวปฏิบัติที่ดีในการนำคอมพิวเตอร์เก่ามาใช้งานใหม่ให้เหมาะสมกับการใช้งานและทำการปรับเปลี่ยนคอมพิวเตอร์และอุปกรณ์ ที่ไม่ต้องการอย่างเหมาะสม

2.1.5) แนวปฏิบัติที่ดีในการบริหารจัดการในหน่วยงาน เช่น การปิดเครื่องในขณะที่ไม่ได้ใช้งาน การกำหนดการใช้ Power Management แทนการใช้ Screen Saver การปรับแสงหน้าจอที่พอเหมาะกับแสงในห้อง การใช้กระดาษให้น้อยลง และพยายามจัดพิมพ์ 2 หน้า และใช้โหมดประหยัดหมึก เป็นต้น

2.2) ซอฟต์แวร์ โดยการมีแนวปฏิบัติที่ดีในการใช้ซอฟต์แวร์ เช่น การติดตั้งซอฟต์แวร์ช่วยประหยัดพลังงาน การติดตั้งที่รองรับการทำงานแบบเสมือน เป็นต้น

2.3) เครือข่ายการสื่อสารพื้นฐาน (Infrastructure) ในหน่วยงานมีการติดตั้งเครือข่ายการสื่อสารของบุคลากรอย่างทั่วถึง เพื่อใช้งาน เช่น การใช้เครื่องพิมพ์ร่วมกัน การใช้ระบบงานสารสนเทศของมหาวิทยาลัย หรือเพื่อการสื่อสารอื่น ๆ เป็นต้น

3) ระดับมหาวิทยาลัย (University Level) ประกอบด้วยส่วนย่อย ๆ ดังนี้

3.1) การใช้เทคโนโลยีแบบเสมือน (Virtualization) มหาวิทยาลัย สามารถกำหนดแนวทางการใช้เทคโนโลยีแบบเสมือน เพื่อลดจำนวนคอมพิวเตอร์เซิร์ฟเวอร์ในศูนย์ข้อมูล

3.2) การใช้ศูนย์ข้อมูล (Data Center) มหาวิทยาลัย สามารถกำหนดแนวทางการใช้ศูนย์ข้อมูล โดยรวมทุกอย่างไว้

ที่แห่งเดียวจึงเป็นการประหยัดพลังงาน และง่ายต่อการควบคุมดูแล

3.3) การใช้ระบบคลาวด์คอมพิวเตอร์ (Cloud Computing) มหาวิทยาลัย สามารถเข้าพื้นที่การบริการต่าง ๆ จากระบบ คลาวด์สาธารณะ ส่งผลให้ลดค่าใช้จ่ายในการดูแลศูนย์ข้อมูลของมหาวิทยาลัย

3.4) การใช้ระบบสารสนเทศเพื่อการบริหาร (Information System) มหาวิทยาลัย สามารถลดการใช้กระดาษ ที่ช่วยลดการเดินทาง ลดการใช้พลังงานของสำนักงาน เนื่องจากอยู่แห่งใดก็สามารถสืบค้นข้อมูล หรือทำงานกับข้อมูลได้

3.5) การทำงานทางไกล/การประชุมทางไกล (Teleworking/conferencing) มหาวิทยาลัยสามารถกำหนดให้ใช้เทคโนโลยีสารสนเทศช่วยการทำงาน ที่บุคลากรไม่จำเป็นต้องอยู่ในหน่วยงาน ซึ่งจะช่วยลดการเดินทาง ลดค่าใช้จ่าย ลดมลพิษทางจราจร ลดพลังงานสำนักงาน

4.4 ผลศึกษาความคิดเห็น ในด้านความเหมาะสมของกรอบการดำเนินงานกรีนไอทีสำหรับมหาวิทยาลัยราชภัฏ มีรายละเอียดดังแสดงตารางที่ 3

ตารางที่ 3 : ผลศึกษาความคิดเห็น

รายการ	สถิติ		การแปลผล
	\bar{X}	SD.	
1. ความเหมาะสมด้านองค์ประกอบ	4.52	0.51	มากที่สุด
1.1 นโยบาย (Policy)	4.60	0.55	มากที่สุด
1.2 ทักษะ (Attitude)	4.40	0.55	มาก
1.3 เทคโนโลยี (Technology)	4.60	0.55	มากที่สุด
1.4 การปฏิบัติการ (Practice)	4.60	0.55	มากที่สุด
1.5 การประเมินผล (Evaluation)	4.40	0.55	มาก
2. ความเหมาะสมด้านระดับในการดำเนินงาน	4.20	0.41	มาก
2.1 ระดับส่วนบุคคล	4.20	0.45	มาก
2.2 ระดับหน่วยงาน	4.00	0.00	มาก
2.3 ระดับมหาวิทยาลัย	4.40	0.55	มาก
เฉลี่ยในภาพรวม	4.40	0.50	มาก

จากตารางที่ 3 ผู้เชี่ยวชาญมีความเห็นโดยรวมในระดับมาก ($\bar{X}=4.40$ และ $SD. = 0.50$) เมื่อพิจารณาองค์ประกอบพบว่า ผู้เชี่ยวชาญมีความเห็นในระดับมากที่สุด ($\bar{X}=4.52$ และ $SD. = 0.51$) และด้านระดับในการดำเนินงาน พบว่า ผู้เชี่ยวชาญมีความเห็นในระดับมาก ($\bar{X}=4.20$ และ $SD. = 0.41$)

5. สรุปผล

จากการวิจัย โดยการสังเคราะห์เอกสารงานวิจัย พบว่า กรอบการดำเนินงานกรีนไอทีสำหรับมหาวิทยาลัยราชภัฏ มี 5 องค์ประกอบ ได้แก่ 1) ด้านนโยบาย (Policy) 2) ด้านทัศนคติ (Attitude) 3) ด้านเทคโนโลยี (Technology) 4) การปฏิบัติการ (Practice) 5) การประเมินผล (Evaluation)

ในส่วนของกรอบการดำเนินงานจัดแบ่งเป็น 3 ระดับ ได้แก่ 1) ระดับส่วนบุคคล (Personal Computing) 2) ระดับหน่วยงาน และ 3) ระดับมหาวิทยาลัย ซึ่งจากผลการศึกษาค้นคว้าของผู้เชี่ยวชาญมีความเห็นในความเหมาะสมในระดับมาก

เอกสารอ้างอิง

- [1] เทวคำปาเชื้อ. เทคโนโลยีสารสนเทศสีเขียว. วารสารเทคโนโลยีสารสนเทศ ปีที่ 5 ฉบับที่ 9 มกราคม - มิถุนายน 2552, หน้า 63-66.
- [2] วิรินทร์ เภมประดิษฐ์สิน. แนวโน้มและเหตุผลในการประยุกต์ใช้ Green IT ที่ควรพิจารณาก่อนที่จะสายเกินไป. วารสารไมโครคอมพิวเตอร์ ปีที่ 29 ฉบับที่ 313 สิงหาคม 2011, หน้า 63-77.
- [3] Center for Sustainable Systems. Green IT. [online] Available at css.snre.umich.edu/css_doc/CSS09-07.pdf. [10/1/2014].
- [4] Chris Dickson. Green IT at Cardiff University – Overview [online] Available at www.cardiff.ac.uk/insrv/chris-dickson [10/1/2014].
- [5] Graem Philipson. Green IT now imperative for today's corporations—Connection Research. [online] Available at <http://www.greentelecomlive.com/2009/10/23/green-it-now-imperative-for-today%E2%80%99s-corporations-%E2%80%93-connection-research/> [10/1/2014].
- [6] Graem Philipson. Green IT Performance Internationally. [online] Available at <http://www.totalexec.com.au/totalexec-views/2010/8/30/latest-research-green-it-performance-internationally.html>. [10/1/2014].
- [7] Green IT Foundation. The smart Green IT Framework. [online] Available at <http://www.g4f.com.br/noticias/detalhado/24> [10/1/2014].
- [8] Open Green IT Policy. Green IT Framework. [online] Available at http://www.opengreenitpolicy.org/?page_id=5 [10/1/2014].
- [9] S. Murugesan, Harnessing. Green IT:Principles and Practices, *IT Pro, IEEE Computer Society*, January/February, 2008. pp 24-33.
- [9] S. Murugesan, Harnessing. Green IT:Principles and Practices, *IT Pro, IEEE Computer Society*, January/February, 2008. pp 24-33.

การประยุกต์ใช้สื่อ eDLTV เพื่อการเรียนรู้สู่การพัฒนาแอปพลิเคชัน เรื่อง รูปทรงเรขาคณิต

The Application of eDLTV Media for Learning to Develop on Geometry

ดวงเนตร ยศคำลือ (Duangnet Yotkamlue)¹ และ วรปภา อารีราษฎร์ (Worapapha Arreerard)²

สาขาคอมพิวเตอร์ศึกษา คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏมหาสารคาม

¹Duang_ne@hotmail.com , ²dr.worapapha@hotmail.com

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ 1) พัฒนาหลักสูตรอบรม การประยุกต์ใช้สื่ออีดีแอลทีวีเพื่อการเรียนรู้สู่การพัฒนา แอปพลิเคชัน เรื่อง รูปทรงเรขาคณิต 2) ศึกษาความรู้และความ เข้าใจของผู้เข้ารับการอบรมตามหลักสูตรอบรมที่พัฒนาขึ้น 3) ศึกษาคุณภาพแอปพลิเคชันของผู้เข้ารับการอบรมตาม หลักสูตรที่พัฒนาขึ้น และ 4) ศึกษาความพึงพอใจของผู้เข้ารับ การอบรมตามหลักสูตรที่พัฒนาขึ้น กลุ่มเป้าหมาย คือ ครูผู้สอนชั้นประถมศึกษาปีที่ 1 กลุ่มสาระการเรียนรู้ คณิตศาสตร์ สำนักงานเขตพื้นที่การศึกษาประถมศึกษา มหาสารคาม เขต 1 จำนวน 30 คน เครื่องมือที่ใช้ในการวิจัย คือ หลักสูตรอบรม แบบประเมินคุณภาพหลักสูตรการอบรม แบบทดสอบวัดผลความรู้และความเข้าใจของผู้เข้ารับ การอบรม แบบประเมินผลการพัฒนาแอปพลิเคชันของผู้เข้ารับ การอบรมและแบบสอบถามความพึงพอใจของผู้เข้ารับ การอบรม สถิติที่ใช้ในการวิเคราะห์ข้อมูล คือ ค่าเฉลี่ย ส่วน เบี่ยงเบนมาตรฐาน และสถิติทดสอบสมมติฐาน *t-test* (Dependent Samples)

ผลการวิจัยพบว่า 1) หลักสูตรอบรมการประยุกต์ใช้สื่ออีดี แอลทีวีเพื่อการเรียนรู้สู่การพัฒนาแอปพลิเคชัน เรื่อง รูปทรง เรขาคณิต ประกอบด้วย 5 ขั้นตอน คือ ขั้นที่ 1 สร้างความตระหนัก ขั้นที่ 2 เลือกเนื้อหาจากกลุ่มสาระการเรียนรู้คณิตศาสตร์ ขั้นที่ 3 ทำ เกมเพื่อการเรียนรู้ ขั้นที่ 4 รวบรวมเครื่องมือ และขั้นที่ 5 ประเมิน แอปพลิเคชันที่พัฒนาขึ้น ผู้เชี่ยวชาญมีความคิดเห็นโดยรวมใน ระดับเหมาะสมมาก ($\bar{X}=4.27, S.D.=0.23$) 2) ผู้เข้ารับการ อบรมมีคะแนนความรู้และความเข้าใจเฉลี่ยหลังการอบรมสูง กว่าก่อนการอบรม อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 3) คุณภาพของแอปพลิเคชันที่ผู้เข้ารับการอบรมพัฒนาขึ้น มีคุณภาพอยู่ในระดับมาก ถึง มากที่สุด ($\bar{X}=4.25-4.83, S.D.=0.39-0.62$) 4) ผู้เข้ารับการอบรมมีความพึงพอใจของต่อ

หลักสูตรที่พัฒนาขึ้นอยู่ในระดับมากที่สุด ($\bar{X}=4.66$, $S.D.=0.19$)

คำสำคัญ : หลักสูตรอบรม สื่ออีดีแอลทีวี แอปพลิเคชัน

Abstract

This research aims to 1) Developing the training syllabus of eDLTV application for learning to developed application on Geometry 2) to study the learning and understanding of the participated teachers in this training program 3) to study the performance of developed application of the participated teachers of this training program and 4) to study the satisfaction of the participated teachers in this program. 30 mathematic teachers of the office of Mahasarakham Primary Education Service Area 1 were selected as the sample group. The research tools were handbooks of training syllabus, the assessment of training syllabus, the test to assess the learning perception of the participated teachers, the assessment for application of participated teachers, and the satisfaction questionnaire for the participated teacher for the syllabus. The statistics used in data analysis were Average, Standard Deviation and *t-test* (Dependent Sample).

The results were found that 1) The training syllabus for the application of eDLTV media for learning to developed application on Geometry consisted of 5 stages: 1 Awareness 2 selected the content from the subject of mathematics 3 provided the games for learning 4) collected the tools and 5) assessed the developed application. The specialists had estimated overview at the high appreciated level ($\bar{X}=4.27, S.D.=0.23$). 2) The participated teachers were acknowledged and understand

after the training syllabus had Average mean score at 0.05 levels statistically significant. 3) The Quality of developed application had high quality level to highest level. and 4) The participated teachers have satisfaction on the developed syllabus at the highest score ($\bar{X}=4.66$, $S.D.=0.19$).

Keywords: Training Syllabus, eDLTV Media, Application

1. บทนำ

ความเจริญทางด้านเทคโนโลยีสารสนเทศและการสื่อสาร หรือ ไอซีที (Information and communication Technology : ICT) ได้มีการพัฒนาและนำมาประยุกต์ใช้อย่างกว้างขวางจากความเจริญก้าวหน้าด้านเทคโนโลยีและการสื่อสารที่ทันสมัย ได้มีการประดิษฐ์ คิดค้นนวัตกรรมใหม่จากคอมพิวเตอร์แบบพกพา หรือ โน้ตบุ๊ก (Note Book) เปลี่ยนเป็นคอมพิวเตอร์พกพา รูปแบบใหม่ เป็นคอมพิวเตอร์แท็บเล็ต (Tablet Computer) ซึ่งเป็นที่นิยมใช้กันอย่างกว้างขวางทั่วโลก จากนโยบายภาครัฐเกี่ยวกับ การจัดเครื่องคอมพิวเตอร์แท็บเล็ตให้แก่โรงเรียนระดับชั้นประถมศึกษาปีที่ 1 ปีการศึกษา 2555 ควบคู่กับการเร่งพัฒนาเนื้อหาที่เหมาะสมตามหลักสูตรบรรจุลงในคอมพิวเตอร์แท็บเล็ต เป็นกลไกในการเปลี่ยนกระบวนการเรียนรู้ โดยส่งเสริมให้นักเรียนทุกระดับชั้นใช้อุปกรณ์คอมพิวเตอร์แท็บเล็ตเพื่อการศึกษา การปฏิบัติในช่วงแรกโครงการ One Tablet PC Per Child หรือ OTPC มุ่งเป้าหมายไปที่กลุ่มผู้ใช้คือนักเรียนชั้นประถมศึกษาปีที่ 1 เป็นกลุ่มนำร่องที่สำคัญของการนำสื่อแท็บเล็ตสู่การพัฒนาการเรียนรู้ในครั้งนี้ [1]

ปีการศึกษา 2556 สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน หรือ สพฐ. ได้จัดทำโครงการจัดการประกวดและอบรมเชิงปฏิบัติการ ผลิตสื่อการเรียนการสอนเพื่อใช้กับแท็บเล็ต โดยสื่อมี 3 รูปแบบ คือ สื่ออิเล็กทรอนิกส์ (Multimedia eBook) การ์ตูนอิเล็กทรอนิกส์ (e-Cartoon) และแอปพลิเคชันการเรียนการสอน (Learning Application) สพฐ. ได้จัดเตรียมเครื่องมือสร้างสื่อและวัตถุดิบ สนับสนุนและกระตุ้น เพื่อให้มีการพัฒนาสื่อเพื่อการเรียนรู้ในรูปแบบแอปพลิเคชันบนระบบปฏิบัติการแอนดรอยด์ที่มีคุณภาพและปริมาณเพียงพอต่อการใช้งาน โดยจัดอบรมกลุ่มครูและบุคลากรทางการศึกษารวมทั้งสิ้น 1,000 คน [2]

มหาวิทยาลัยราชภัฏมหาสารคาม [4] เป็นสถาบันการศึกษาเพื่อพัฒนาท้องถิ่น ได้ดำเนินการส่งเสริมการเรียนรู้ของชุมชน โดยใช้ไอซีทีในภาคตะวันออกเฉียงเหนือ ภายใต้การดำเนินงานของโครงการศูนย์ทางไกลเพื่อพัฒนาการศึกษาและพัฒนาชนบทในภาคตะวันออกเฉียงเหนือ ที่ดำเนินการโดยคณะเทคโนโลยีสารสนเทศ ร่วมกับเครือข่ายทั้งหน่วยงานภาครัฐและเอกชน การส่งเสริมการใช้สื่ออีดีแอลทีวี (eDLTV : Electronic Distance Learning Via Television) เป็นโครงการหนึ่งในการดำเนินงาน ทั้งนี้สื่ออีดีแอลทีวีเป็นสื่อที่ผลิตขึ้นภายใต้โครงการจัดทำเนื้อหาบนระบบอีเลิร์นนิ่ง (eDL-Square) [5] ซึ่งเป็นโครงการหนึ่งของโครงการเทคโนโลยีสารสนเทศตามพระราชดำริ สมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี เพื่อให้ครูได้ใช้เป็นสื่อในการเรียนการสอน ซึ่งสื่ออีดีแอลทีวี เป็นสื่อที่มีเนื้อหาสาระครบทุกระดับชั้น จากการสำรวจโรงเรียนและความต้องการของครูในการใช้สื่ออีดีแอลทีวีเพื่อการเรียนการสอน [3] พบว่า ครูและบุคลากรทางการศึกษามีความสนใจในการนำสื่อ eDLTV ไปใช้ในการจัดการเรียนรู้ร้อยละ 92.44 โดยนำไปเป็นแหล่งสืบค้น แหล่งเรียนรู้ ร้อยละ 88.02 และเป็นสื่อเสริม ร้อยละ 82.23 และ ครูและบุคลากรทางการศึกษามีความสนใจในการพัฒนาต่อขยายการจัดการเรียนรู้ด้วยสื่ออิเล็กทรอนิกส์ ร้อยละ 86.23 จากรายงานผลการอบรมครูพัฒนาแอปพลิเคชันที่ สพฐ. ได้จัดขึ้น โดยมีครูที่เข้ารับการอบรมเพียง 1,000 คน ยังไม่ครอบคลุมครูที่สังกัด สพฐ. เพื่อเป็นการพัฒนาครูในการสร้างแอปพลิเคชันเพื่อการเรียนรู้อำนาจงานเขตพื้นที่การศึกษามหาสารคามเขต 1 ได้ร่วมกับมหาวิทยาลัยราชภัฏมหาสารคาม เพื่อให้มีการพัฒนาครูและบุคลากรทางการศึกษาส่งเสริมในการใช้ คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนการสอน และการประยุกต์ใช้สื่ออีดีแอลทีวีเพื่อการเรียนการสอน [4]

จากหลักการและเหตุผลที่กล่าวมา ผู้วิจัยจึงสนใจที่จะศึกษาการประยุกต์ใช้สื่ออีดีแอลทีวีเพื่อการเรียนรู้สู่การพัฒนาแอปพลิเคชัน เรื่อง รูปทรงเรขาคณิต โดยมุ่งอบรมครูชั้นประถมศึกษาปีที่ 1 เนื่องจากเป็นกลุ่มเป้าหมายที่จะต้องจัดการเรียนการสอนโดยใช้แท็บเล็ตตามนโยบายภาครัฐ การวิจัยครั้งนี้จึงเป็นการจัดทำหลักสูตรอบรมครูเพื่อศึกษาเนื้อหา สาระการเรียนรู้จากสื่ออีดีแอลทีวี นำมาสู่การออกแบบแอปพลิเคชันเพื่อ

การเรียนรู้ให้ครูผู้สอนสามารถสร้างแอปพลิเคชันบนแท็บเล็ตให้กับนักเรียนชั้นประถมศึกษาปีที่ 1 กลุ่มสาระการเรียนรู้คณิตศาสตร์ เรื่อง รูปทรงเรขาคณิต ที่มีทั้งเนื้อหาสาระ เกม และแบบทดสอบ บรรจุไว้ในคอมพิวเตอร์แท็บเล็ต เพื่อเป็นการส่งเสริมการดำเนินงานของมหาวิทยาลัยที่ ให้บริการส่งเสริมการพัฒนาครูในเขตพื้นที่บริการสอดคล้อง กับนโยบายภาครัฐต่อไป

2. วัตถุประสงค์

- 2.1. เพื่อพัฒนาหลักสูตรอบรมการประยุกต์ใช้สื่ออีดีแอลทีวีเพื่อการเรียนรู้สู่การพัฒนาแอปพลิเคชัน เรื่อง รูปทรงเรขาคณิต
- 2.2. เพื่อศึกษาความรู้และความเข้าใจของผู้เข้ารับการอบรมตามหลักสูตรอบรมที่พัฒนาขึ้น
- 2.3. เพื่อศึกษาคุณภาพแอปพลิเคชันของผู้เข้ารับการอบรมตามหลักสูตรที่พัฒนาขึ้น
- 2.4. เพื่อศึกษาความพึงพอใจของผู้เข้ารับการอบรมตามหลักสูตรที่พัฒนาขึ้น

3. วรรณกรรมที่เกี่ยวข้อง

3.1 สื่ออีดีแอลทีวี

อีดีแอลทีวี เป็นสื่อที่นำเอาเนื้อหาที่ออกอากาศทางสถานีวิทยุและโทรทัศน์การศึกษาทางไกลผ่านดาวเทียม จากโรงเรียนวังไกลกังวล จังหวัดประจวบคีรีขันธ์ หรือ การศึกษาทางไกลผ่านดาวเทียม (DLTV) ที่ประกอบด้วย คู่มือครู วัสดุทัศนการสอน ใบความรู้ ใบงาน และแบบทดสอบ นำเสนอเนื้อหาบนระบบ e-Learning ของการศึกษาทางไกลผ่านดาวเทียม ให้บริการเผยแพร่ในรูปแบบ Off-line และ On-line ผ่านทางอินเทอร์เน็ตให้แก่ครู นักเรียน และผู้สนใจทั่วไปได้ใช้ประโยชน์ในการเรียนการสอน หรือศึกษาเพิ่มเติมโดยสำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ(สวทช.) กำกับดูแลการเผยแพร่ [5]

3.2 การพัฒนาหลักสูตรอบรม

ADDIE Model เป็นรูปแบบที่ได้รับการยอมรับอย่างกว้างขวางในการนำมาพัฒนาหลักสูตรอบรม มีขั้นตอนในการออกแบบพัฒนาหลักสูตร แสดงดังแผนภาพที่ 1



ภาพที่ 1: ขั้นตอนการพัฒนาหลักสูตรตาม ADDIE Model

จากภาพที่ 1 ADDIE Model ประกอบด้วยขั้นตอนทั้งหมด 5 ขั้นตอน ได้แก่ 1) ขั้นตอนการวิเคราะห์(Analysis) เป็นขั้นวางแผนวิเคราะห์ปัญหาหรือเตรียมการสื่อต่าง ๆ ที่จำเป็นต่อการพัฒนาหลักสูตร 2) ขั้นตอนการออกแบบ (Design) เป็นขั้นตอนการนำผลลัพธ์จากขั้นตอนการวิเคราะห์ เพื่อวางแผนสำหรับพัฒนาหลักสูตร 3) ขั้นตอนการพัฒนา (Development) เป็นขั้นตอนการพัฒนาสร้างขึ้นบนขั้นตอนการวิเคราะห์และการออกแบบ 4) ขั้นตอนการทดลองใช้ (Implementation) เป็นขั้นที่นำหลักสูตรที่มีองค์ประกอบครบสมบูรณ์มาทดลองใช้เพื่อหาประสิทธิภาพของหลักสูตรการอบรม และ 5) ขั้นตอนการประเมินผล (Evaluation) เป็นขั้นตอนสุดท้าย โดยการนำผลการทดลองที่ได้มาสรุปหาประสิทธิภาพของหลักสูตร ซึ่งกระบวนการทั้ง 5 ขั้นทำให้ได้หลักสูตรครบถ้วนเหมาะสมกับการอบรมได้เป็นอย่างดี [6]

3.3 งานวิจัยที่เกี่ยวข้อง

อรฉันทน์ เพื่อนสิ่ง ได้ทำวิจัย เรื่อง การพัฒนาหลักสูตรอบรมการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ เรื่อง การทำภาพสติกเกอร์ด้วย Ipad2 สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 1 พบว่า 1) หลักสูตรอบรมประกอบด้วยหลักสูตรอบรม เนื้อหาความรู้พื้นฐานเกี่ยวกับแท็บเล็ต การใช้งานบนกล้อง Ipad2 และการทำภาพสติกเกอร์ด้วย Ipad2 ผู้เชี่ยวชาญมีความคิดเห็นโดยรวมต่อหลักสูตรอบรมอยู่ในระดับเหมาะสมมาก 2) ผลการเปรียบเทียบผลการเรียนรู้ก่อนอบรมและหลังอบรมของผู้เข้ารับการอบรมด้วยหลักสูตรที่พัฒนาขึ้น ผู้เข้ารับการอบรมมีความรู้ความเข้าใจในการอบรม โดยมีคะแนนหลังอบรมสูงกว่าก่อนอบรมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และ 3) ความพึงพอใจของผู้เข้ารับการอบรมด้วยหลักสูตรที่พัฒนาขึ้นเฉลี่ยโดยรวมอยู่ในระดับมากที่สุด [7]

นงเยาว์ ประธิปอาราม ทำวิจัยเรื่อง การพัฒนาหลักสูตรอบรมการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ เรื่อง การตกแต่งรูปภาพโดยใช้ Ps Express ด้วย iPad 2 สำหรับชั้นมัธยมศึกษาปีที่ 4 พบว่า 1) หลักสูตรอบรมที่พัฒนาขึ้นมีคุณภาพ

ในระดับมาก 2) ผลการเปรียบเทียบคะแนนหลังการอบรมสูงกว่าอบรมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และ 3) ผู้เข้ารับการอบรมมีความพึงพอใจต่อหลักสูตรการอบรมในระดับมาก [8]

หมกนุช โคตรธาดา ได้ศึกษาวิจัยเรื่อง การพัฒนาหลักสูตรอบรมการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ เรื่อง การสร้างงานนำเสนอด้วย iPad 2 สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 5 พบว่า 1) หลักสูตรอบรมประกอบด้วย หลักสูตรอบรมเนื้อหาเกี่ยวกับแท็บเล็ตเบื้องต้น การใช้งานแอปพลิเคชันบน Ipad2 และการสร้างงานนำเสนอด้วย iPad 2 ผลการประเมินคุณภาพหลักสูตรอยู่ในระดับเหมาะสมมาก 2) ผลการเปรียบเทียบผลการเรียนรู้ก่อนอบรมและหลังอบรมของผู้เข้ารับการอบรม มีคะแนนหลังอบรมสูงกว่าก่อนอบรมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และ 3) ความพึงพอใจของผู้เข้ารับการอบรมด้วยหลักสูตรอบรมที่พัฒนาขึ้น เฉลี่ยโดยภาพรวมอยู่ในระดับมาก [9]

Nicholas Gorgievski and others.[2005]ได้ทำการศึกษาเรื่อง เครื่องมือสำหรับการเรียนการสอนแคลคูลัส พบว่า การรับรู้ของนักเรียนจากเครื่อง Tablet PC เป็นเครื่องมือในการเรียนการสอนสำหรับการสอนแคลคูลัส ซึ่งวัตถุประสงค์ของการสำรวจครั้งนี้ คือ การเก็บรวบรวมข้อมูลเกี่ยวกับการรับรู้ของนักเรียนเกี่ยวกับการใช้งานของเครื่อง Tablet PC เป็นเครื่องมือในการเรียนการสอนได้อย่างมีประสิทธิภาพและมีประสิทธิภาพครอบคลุมวัสดุการเรียนการสอนในชั้นเรียนผลการชี้ให้เห็นว่านักเรียนรับรู้ว่าการใช้ Tablet PC มีความสนใจในการเรียนในชั้นเรียน ช่วยให้เข้าใจเนื้อหาในเรื่องที่เรียนครอบคลุมเนื้อหาในการเรียน ทำให้เกิดประสิทธิภาพในการเรียน [10]

4. วิธีการดำเนินการวิจัย

4.1 ขั้นตอนการวิจัย ผู้วิจัยดำเนินการตามรูปแบบ ADDIE ดังนี้

4.1.1 ขั้นการวิเคราะห์ ดำเนินการศึกษาเอกสารงานวิจัยที่เกี่ยวข้อง ศึกษาในด้านการพัฒนาหลักสูตร การประยุกต์ใช้แท็บเล็ตเพื่อการเรียนรู้ การพัฒนาแอปพลิเคชัน

4.1.2 ขั้นการออกแบบ ดำเนินการออกแบบหลักสูตรอบรม การประยุกต์ใช้สื่อ eDLTV เพื่อการเรียนรู้สู่การพัฒนา

แอปพลิเคชัน เรื่อง รูปทรงเรขาคณิต และออกแบบเครื่องมือการประเมินผลต่าง ๆ

4.1.3 ขั้นการพัฒนา ดำเนินการพัฒนาหลักสูตรอบรมการประยุกต์ใช้สื่อ eDLTV เพื่อการเรียนรู้สู่การพัฒนาแอปพลิเคชัน เรื่อง รูปทรงเรขาคณิตและเครื่องมือการประเมินผลต่าง ๆ ตลอดจนหาคุณภาพของหลักสูตร และเครื่องมือต่าง ๆ

4.1.4 ขั้นทดลองใช้ ดำเนินการทดลองใช้กับกลุ่ม เป้าหมายตามแบบแผนการวิจัย จัดเก็บข้อมูลโดยใช้เครื่องมือการวิจัย

4.1.5 ขั้นการประเมินผล ดำเนินการวิเคราะห์ และสรุปผลการวิจัย

4.2 กลุ่มเป้าหมายการวิจัย คือ ครูผู้สอนชั้นประถมศึกษาปีที่ 1 กลุ่มสาระการเรียนรู้คณิตศาสตร์ สำนักงานเขตพื้นที่การศึกษาประถมศึกษามหาสารคาม เขต 1 จำนวน 30 คน ที่สมัครใจเข้ารับการอบรม

4.3 ตัวแปรการวิจัย ประกอบด้วย

ตัวแปรต้น ได้แก่ หลักสูตรอบรมการประยุกต์ใช้สื่อ eDLTV เพื่อการเรียนรู้สู่การพัฒนาแอปพลิเคชัน เรื่อง รูปทรงเรขาคณิต

ตัวแปรตาม ได้แก่ คุณภาพของหลักสูตรอบรมที่พัฒนาขึ้น ความรู้ความเข้าใจของผู้รับการอบรม คุณภาพของแอปพลิเคชันที่ผู้เข้ารับการอบรมพัฒนาขึ้น และความพึงพอใจของผู้เข้ารับการอบรม

4.4 เครื่องมือการวิจัย ประกอบด้วย

4.4.1 หลักสูตร

4.4.2 แบบประเมินหลักสูตร

4.4.3 แบบทดสอบวัดผลความรู้และความเข้าใจ

4.4.4 แบบประเมินคุณภาพแอปพลิเคชัน

4.4.5 แบบประเมินความพึงพอใจ

4.5 สถิติที่ใช้วิเคราะห์ข้อมูล ได้แก่ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และสถิติ t-test แบบ (Dependent Sample)

5. ผลการวิจัย

5.1 ผลการพัฒนาหลักสูตรอบรม

ผู้วิจัยได้พัฒนาหลักสูตรอบรมการประยุกต์ใช้สื่ออีดีแอลทีวี เพื่อการเรียนรู้สู่การพัฒนาแอปพลิเคชัน เรื่อง รูปทรงเรขาคณิต ประกอบด้วยกระบวนการอบรม 5 ขั้นตอน ดังนี้

ขั้นที่ 1 สร้างความตระหนัก : เป็นการกล่าวถึงนโยบายการแจกแท็บเล็ตเพื่อการศึกษา และนำการใช้งานสื่ออีดีแอลที่วีการใช้งานแอปพลิเคชันเพื่อการเรียนรู้ ตัวอย่างแอปพลิเคชันเพื่อการเรียนรู้บนแท็บเล็ต และแนะนำการใช้งานเว็บไซต์ www.otpcappcon.com

ขั้นที่ 2 เลือกเนื้อหาหลักสูตรการเรียนรู้คณิตศาสตร์ : ผู้เข้ารับการอบรมเลือกเนื้อหาเกี่ยวกับ รูปทรงเรขาคณิต นำเสนอในรูปแบบต่างๆ ที่สอดคล้องกับกิจกรรมการเรียนการสอน และนำสู่การพัฒนาเนื้อหาเพื่อพัฒนาแอปพลิเคชัน

ขั้นที่ 3 ทำเกมเพื่อการเรียนรู้ : จัดทำเกมจากเนื้อหาที่เลือกประกอบด้วย 5 เกม ได้แก่ เกมจับคู่ เกมลากเส้นจับคู่ภาพที่เหมือนกัน เกมค้นหาภาพ เกมแยกประเภท และเกมเลือกภาพ

ขั้นที่ 4 รวบรวมเครื่องมือ : เป็นขั้นตอนการนำเนื้อหาและเกมที่สร้างขึ้นมาจัดทำเป็นแอปพลิเคชันและตรวจสอบความถูกต้อง ปรับปรุงแก้ไขและนำขึ้นบนแท็บเล็ต

ขั้นที่ 5 ประเมินแอปพลิเคชันที่พัฒนาขึ้น : การประเมินแอปพลิเคชันที่พัฒนาขึ้น โดยวิทยากรและผู้ช่วยวิทยากร รวม 5 คน โดยใช้แบบประเมินที่ผู้วิจัยพัฒนาขึ้น

ทั้งนี้พิจารณาความเหมาะสม โดยใช้แบบประเมินคุณภาพหลักสูตรอบรม โดยผู้เชี่ยวชาญจำนวน 5 คน ผลการประเมินแสดงในตารางที่ 1

ตารางที่ 1: ผลการประเมินคุณภาพหลักสูตรอบรม

รายการประเมิน	ผลการประเมิน		
	\bar{X}	S.D.	ความหมาย
1) หลักการและเหตุผล	4.20	0.29	มาก
2) วัตถุประสงค์ของหลักสูตรอบรม	4.20	0.44	มาก
3) โครงสร้างของหลักสูตรอบรม	4.30	0.20	มาก
4) ด้านการจัดการ/กิจกรรมการอบรม	4.10	0.22	มาก
5) การวัดผลการอบรม	4.30	0.57	มาก
6) การประเมินผลการอบรม	4.20	0.57	มาก
ความเหมาะสมเฉลี่ยโดยรวม	4.27	0.23	มาก

จากตารางที่ 1 พบว่า หลักสูตรอบรมการประยุกต์ใช้สื่อ eDLTV เพื่อการเรียนรู้สู่การพัฒนาแอปพลิเคชัน เรื่อง รูปทรงเรขาคณิต ผู้เชี่ยวชาญมีความคิดเห็นโดยรวมในระดับเหมาะสมมาก ($\bar{X}=4.27, S.D.=0.23$)

4.2 ผลการศึกษาความรู้ความเข้าใจของผู้เข้ารับการอบรม

ผู้วิจัยทดสอบความรู้และความเข้าใจของผู้เข้ารับการอบรมจำนวน 30 คน ทั้งก่อนอบรมและหลังการอบรมและนำคะแนนที่ได้ มาวิเคราะห์ค่าจำนวนค่าสถิติ t-test (Dependent Samples) ผลการวิเคราะห์แสดงในตารางที่ 2

ตารางที่ 2: ผลการวิเคราะห์ความรู้และความเข้าใจของผู้เข้ารับการอบรม

คะแนนการทดสอบ	ผู้เข้ารับการอบรม	\bar{X}	S.D.	df	t
ก่อนอบรม	30	14.00	4.14	29	27.37
หลังอบรม	30	29.27	5.36		

จากตารางที่ 2 พบว่า ค่า t ที่คำนวณได้มีค่า 27.37 ซึ่งมีค่ามากกว่าค่า $t_{(ตาราง \alpha=.05, df=29)}$ มีค่าเท่ากับ 1.699 สรุปได้ว่า คะแนนความรู้และความเข้าใจเฉลี่ยหลังการอบรมสูงกว่าก่อนการอบรมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

4.3 ผลการศึกษาคุณภาพแอปพลิเคชันของผู้เข้ารับการอบรม

ผู้วิจัยนำแอปพลิเคชันจากผู้เข้ารับการอบรมหลักสูตรอบรมการประยุกต์ใช้สื่อ eDLTV เพื่อการเรียนรู้สู่การพัฒนาแอปพลิเคชัน เรื่อง รูปทรงเรขาคณิตที่พัฒนาขึ้น จำนวน 30 ชิ้นงาน ประเมินคุณภาพโดยวิทยากรและผู้ช่วยวิทยากร รวม 5 คน ผลการประเมินคุณภาพแอปพลิเคชัน พบว่า แอปพลิเคชันที่ผู้เข้ารับการอบรมพัฒนาขึ้น มีคุณภาพอยู่ในระดับมาก ถึงมากที่สุด ($\bar{X}=4.25-4.83, S.D.=0.39-0.62$)

4.4 ผลการศึกษาความพึงพอใจของผู้เข้ารับการอบรม

ผู้วิจัยได้ศึกษาความพึงพอใจของผู้เข้ารับการอบรมหลังจากได้จัดการอบรมหลักสูตรการประยุกต์ใช้สื่อ eDLTV เพื่อการเรียนรู้สู่การพัฒนาแอปพลิเคชัน เรื่อง รูปทรงเรขาคณิต ที่พัฒนาขึ้น ผลการประเมินแสดงในตารางที่ 3

ตารางที่ 3: ผลการศึกษาความพึงพอใจของผู้เข้ารับการอบรม

รายการประเมิน	ผลการประเมิน		
	\bar{X}	S.D.	ความหมาย
1) ด้านวิทยากร	4.60	0.24	มากที่สุด
2) ด้านกระบวนการอบรม	4.65	0.29	มากที่สุด
3) ด้านความรู้ความเข้าใจ	4.85	0.21	มากที่สุด
4) ด้านการนำไปใช้ประโยชน์	4.83	0.30	มากที่สุด
5) ด้านสถานที่ระยะเวลา	4.36	0.39	มาก
ความพึงพอใจเฉลี่ยโดยรวม	4.66	0.19	มากที่สุด

จากตารางที่ 3 พบว่า ผู้เข้ารับการอบรมมีความพึงพอใจต่อการจัดการอบรมด้วยหลักสูตรการประยุกต์ใช้สื่อ eDLTV เพื่อการเรียนรู้สู่การพัฒนาแอปพลิเคชัน เรื่อง รูปทรงเรขาคณิตที่พัฒนาขึ้น โดยภาพรวมเฉลี่ยทั้ง 5 ด้าน พบว่าผู้เข้ารับการอบรมส่วนใหญ่ มีความพึงพอใจอยู่ในระดับมากที่สุด ($\bar{X}=4.66, S.D.=0.19$)

5. สรุปผลการวิจัย

5.1 ผลการพัฒนาหลักสูตรอบรมการประยุกต์ใช้สื่อ eDLTV เพื่อการเรียนรู้สู่การพัฒนาแอปพลิเคชัน เรื่อง รูปทรงเรขาคณิต ประกอบด้วยกระบวนการอบรม 5 ขั้นตอน ดังนี้ ขั้นที่ 1 สร้างความตระหนัก ขั้นที่ 2 เลือกเนื้อหาในกลุ่มสาระการเรียนรู้คณิตศาสตร์ ขั้นที่ 3 ทำเกมเพื่อการเรียนรู้ ขั้นที่ 4 รวบรวมเครื่องมือ ขั้นที่ 5 ประเมินแอปพลิเคชันที่พัฒนาขึ้น ทั้งนี้ผู้เชี่ยวชาญ 5 ท่าน มีความคิดเห็นโดยรวมในระดับเหมาะสมมาก

5.2 ผลการศึกษาความรู้และความเข้าใจก่อนอบรมและหลังอบรมของผู้เข้ารับการอบรม โดยใช้การทดสอบแบบ t-test (Dependent Samples) สรุปได้ว่า คะแนนความรู้และความเข้าใจเฉลี่ยหลังการอบรมสูงกว่าก่อนการอบรม อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

5.3 ผลการศึกษาคุณภาพแอปพลิเคชันของผู้เข้ารับการอบรม พบว่า แอปพลิเคชันที่ผู้เข้ารับการอบรมพัฒนาขึ้น มีคุณภาพอยู่ในระดับมาก ถึง มากที่สุด

5.4 ผลการศึกษาความพึงพอใจของผู้เข้ารับการอบรม ที่มีต่อหลักสูตรที่พัฒนาขึ้น พบว่า โดยรวมมีความพึงพอใจอยู่ในระดับมากที่สุด

6. เอกสารอ้างอิง

- [1] สำนักงานเลขาธิการรัฐมนตรี, “โครงการแท็บเล็ตเพื่อการศึกษา”. สืบค้นเมื่อ จาก <http://suksa.2moe.go.th> (เมื่อ 10 มกราคม 2556)
- [2] สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน .คู่มือการอบรมการสร้างแอปพลิเคชัน , 2555.
- [3] สมเจตน์ ภูศิริ และ วรปภา อธิราชบุรี, “รายงานการดำเนินงานเครือข่ายเผยแพร่ถ่ายทอดและพัฒนาสื่อการเรียนการสอนบนระบบ e-Learning (eDL-Square)”. มหาวิทยาลัยราชภัฏมหาสารคาม, 2555.
- [4] ราชภัฏมหาสารคาม, มหาวิทยาลัย, “รายงานผลการดำเนินงานส่งเสริมการพัฒนาครูบุคลากรทางการศึกษา และนักเรียนในการใช้เทคโนโลยีสารสนเทศเพื่อยกระดับการศึกษาเรียนรู้ของชุมชนมหาวิทยาลัยราชภัฏมหาสารคาม” มหาสารคาม, : มหาวิทยาลัยราชภัฏมหาสารคาม, 2555.
- [5] โครงการ จัดทำเนื้อหา ระบบ e-Learning ของการศึกษาทางไกลผ่านดาวเทียม เฉลิมพระเกียรติเนื่องในโอกาสสมหามงคล เฉลิมพระชนมพรรษา ๘๐ พรรษา ๕ ธันวาคม ๒๕๕๐. (โครงการ eDLTV) [ออนไลน์].สืบค้นจาก <http://edltv.thai.net> (เมื่อ 12 มีนาคม 2556)
- [6] พิสุทธิ อธิราชบุรี, “การพัฒนาซอฟต์แวร์ทางการศึกษา” มหาสารคาม : มหาวิทยาลัยราชภัฏมหาสารคาม, 2552.
- [7] อรฉัตร เตือนพิ่ง, “การพัฒนาหลักสูตรอบรมการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ เรื่อง การทำภาพสติกเกอร์ ด้วย iPad 2,” มหาวิทยาลัยราชภัฏมหาสารคาม, 2555.
- [8] นงเยาว์ ประธิปอาราม, “การพัฒนาหลักสูตรอบรมการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ เรื่อง การตกแต่งรูปภาพโดยใช้ Ps Express ด้วย iPad 2,” มหาวิทยาลัยราชภัฏมหาสารคาม, 2555.
- [9] ชมภูษ ชโคตรธาดา. “การพัฒนาหลักสูตรอบรมการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ เรื่อง การสร้างงานนำเสนอด้วย iPad2 สำหรับนักเรียนชั้นประถมศึกษาปีที่ 5” : มหาวิทยาลัยราชภัฏมหาสารคาม, 2555.
- [10] Nicholas Gorgievski and other. “Tablet PC: A Preliminary Report on a Tool for Teaching Calculus”. *The International Journal for Technology in Mathematics Education*, 95-102, 2005.

ผลการสังเคราะห์การส่งเสริมครูพัฒนาแอปพลิเคชันเพื่อการเรียนรู้ บนคอมพิวเตอร์แท็บเล็ต ด้วยเทคนิคเพื่อนคู่คิด

Results of the Synthesis of Learning Application Development with Tablet for Teachers by Using Think-Pair-Share Techniqe

วิญญู อุดระ (Winyoo Uttara)¹ วรปภา อารีราษฎร์ (Worapapha Arreerard)² สุรัตน์ ดวงชาตม (Surat Duangchatom)³

¹สาขาคอมพิวเตอร์ศึกษา มหาวิทยาลัยราชภัฏมหาสารคาม

²คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏมหาสารคาม

³ผู้อำนวยการสำนักงานเขตพื้นที่การศึกษาประถมศึกษามหาสารคาม เขต 1

¹winyuu@gmail.com, ²dr.worapapha@hotmail.com, ³surat.d@hotmail.com

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อ 1) สังเคราะห์องค์ประกอบของการส่งเสริมครูพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ต ด้วยเทคนิคเพื่อนคู่คิด และ 2) สังเคราะห์ขั้นตอนการส่งเสริมครูพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ต ด้วยเทคนิคเพื่อนคู่คิด กลุ่มเป้าหมายในการวิจัยเป็นผู้เชี่ยวชาญ จำนวน 12 คน จาก 4 กลุ่ม คือ 1)กลุ่มนักเทคโนโลยีสารสนเทศ 2) กลุ่มผู้บริหารโรงเรียนและศึกษานิเทศก์ 3) กลุ่มวิทยากรด้านการพัฒนาแอปพลิเคชัน และ 4) กลุ่มนักวิชาการศึกษาจากมหาวิทยาลัย โดยใช้การประชุมกลุ่มย่อย เครื่องมือที่ใช้ในการวิจัย คือ แบบสอบถามความคิดเห็นของผู้เชี่ยวชาญที่มีต่อองค์ประกอบและขั้นตอนของการส่งเสริมครูพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ต ด้วยเทคนิคเพื่อนคู่คิด สถิติที่ใช้ในการวิจัย คือ ค่าเฉลี่ย (\bar{X}) และส่วนเบี่ยงเบนมาตรฐาน (S.D.)

ผลการวิจัย พบว่า 1) องค์ประกอบของการส่งเสริมครูพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ต ด้วยเทคนิคเพื่อนคู่คิด ประกอบด้วย 4 ส่วน คือ 1.1) นโยบายแนวคิดและทฤษฎี 1.2) ระบบเทคโนโลยีสารสนเทศและการสื่อสาร 1.3) หน่วยงานสนับสนุน และ 1.4) กระบวนการส่งเสริมและพัฒนาครู โดยความคิดเห็นของผู้เชี่ยวชาญที่มีต่อองค์ประกอบ การส่งเสริมครูพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ต ด้วยเทคนิคเพื่อนคู่คิด สรุปโดยรวมอยู่ในระดับมากที่สุด ($\bar{X} = 4.75$, S.D. = 0.43) และ 2) ผลการสังเคราะห์

ขั้นตอนการส่งเสริมครูพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ต ด้วยเทคนิคเพื่อนคู่คิด มี 4 ขั้นตอน คือ 2.1) ขั้นโน้มน้าวและไตร่ตรอง 2.2) ขั้นฝึกปฏิบัติเพื่อสร้างประสบการณ์ และ 2.3) ขั้นติดตามและสะท้อนผล โดยความคิดเห็นของผู้เชี่ยวชาญที่มีต่อขั้นตอนการส่งเสริมครูพัฒนาแอปพลิเคชันเพื่อการเรียนรู้ บนคอมพิวเตอร์แท็บเล็ต ด้วยเทคนิคเพื่อนคู่คิด สรุปโดยรวมอยู่ในระดับมากที่สุด ($\bar{X} = 4.80$, S.D. = 0.38)

คำสำคัญ : การส่งเสริมครู แอปพลิเคชันเพื่อการเรียนรู้ คอมพิวเตอร์แท็บเล็ต เทคนิคเพื่อนคู่คิด

Abstract

The research aimed at 1) synthesizing components, 2) steps for learning application development with tablet for teachers by using think-pair-share technique. The target population was twelve participants consisting of 4 major experts in 1) information technology, 2) director and supervisor, 3) Application trainers and 4) academicians. Focus group method was employed for the research. The instrument was a questionnaire. The statistics used were mean and standard deviation.

The research showed that 1) components of learning application development with tablet for teachers by using think-pair-share technique consisted of 4 parts: 1.1) policy, theory and concept, 1.2) system of information technology and communication, 1.3) supporting organization,

and 1.4) process of teacher development. In addition, the find showed that the average level of the opinion of the experts towards components and sequence of learning application development with tablet for teachers by using think-pair-share technique was very high ($\bar{X} = 4.75$, $S.D. = 0.43$). 2) The research showed that steps for learning application development with tablet for teachers by using think-pair-share technique consisted of 4 steps: 2.1) practical practice, 2.2) following up and reflection, 2.3) evaluation and feedback. The finding showed that the average level of the opinion of the experts towards the learning application development with tablet for teachers was very high ($\bar{X} = 4.80$, $S.D. = 0.38$).

Keywords: Development for Teachers, Learning Application, Tablet Computer, Think-Pair-Share Technique.

1. บทนำ

เทคโนโลยีสารสนเทศและการสื่อสาร หรือไอซีที มีบทบาทสำคัญในการพัฒนาทรัพยากรมนุษย์ทั้งเชิงปริมาณและคุณภาพ ก่อเกิดประโยชน์ต่อการเรียนรู้ การทำงานและการดำรงชีวิตประจำวันและกรอบนโยบายเทคโนโลยีสารสนเทศและการสื่อสาร ระยะพ.ศ. 2554-2563 ของประเทศไทย (IT 2020) มีเป้าหมายหลักด้านการพัฒนากำลังคนและสังคม กำหนดให้มีโครงสร้างพื้นฐานไอซีทีความเร็วสูง (Broadband) ที่กระจายอย่างทั่วถึงประชาชนสามารถเข้าถึงได้อย่างเท่าเทียมกัน โดยเฉพาะบริการด้านการศึกษา ซึ่งกำหนดให้สถาบันการศึกษาต้องนำไอซีทีมาใช้เป็นเครื่องมือในการเรียนการสอนเพิ่มมากขึ้น ดังนั้นคอมพิวเตอร์แท็บเล็ตจึงกำหนดเป็นนโยบายของรัฐที่สำคัญในฐานะสื่อประกอบการเรียนรู้ที่ช่วยเพิ่มประสิทธิภาพการเรียนรู้ของผู้เรียนให้ก้าวทันสังคมยุคปัจจุบัน [1] ซึ่งสอดคล้องกับงานวิจัยของ Xiang and Colleagues [2] และ El-Gayar and Colleagues [3] ที่ว่าการใช้สื่อแท็บเล็ตเพื่อเป็นเครื่องมือที่ทรงประสิทธิภาพในการเรียนรู้หรือสำหรับการสอนนั้นควรมีการพัฒนาาร่วมกันระหว่างสถานศึกษากับผู้เรียนจากภายนอกสถานศึกษา มีการพัฒนาและสร้างระบบปฏิสัมพันธ์ (Interactive) ของการใช้สื่อระหว่างครูกับนักเรียน มีการสร้างความสมบูรณ์และความชัดเจนในบทเรียน และมีการพัฒนาปรับปรุงรูปแบบวิธีการเรียนรู้อย่างต่อเนื่อง และผู้เรียนส่วนใหญ่มีความต้องการให้โรงเรียน

หรือสถาบันการศึกษาจัดรูปแบบโปรแกรมการใช้สื่อคอมพิวเตอร์แท็บเล็ตที่สนองต่อการสร้างสรรค์งานในสภาพแวดล้อมหรือสภาพการบริหารจัดการที่เหมาะสม รวมทั้งสามารถเอื้ออำนวยประโยชน์ต่อการใช้อุปกรณ์เรียนประเภทคอมพิวเตอร์แท็บเล็ตให้บังเกิดประสิทธิภาพสูงสุด

นโยบายภาครัฐเกี่ยวกับแท็บเล็ตเพื่อการศึกษา (Tablet for Education) ที่มอบคอมพิวเตอร์แท็บเล็ต (Tablet Computer) ประจำตัวนักเรียนเพื่อเป็นเครื่องมือสำหรับการเรียนยุคใหม่ภายใต้โครงการ OTPC “One Tablet PC per Child” กำหนดให้มีการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ของนักเรียนชั้นประถมศึกษาปีที่ 1 ภาคเรียน 1 ปีการศึกษา 2555 และขยายไปสู่ชั้นมัธยมศึกษาปีที่ 1 ในภาคเรียนที่ 1 ปีการศึกษา 2556 ซึ่งในปีการศึกษา 2555 สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (สพฐ.) ได้จัดโครงการประกวดสร้างสื่อการเรียนรู้สู่แท็บเล็ต โดยส่งเสริมให้ครูบุคลากรทางการศึกษา บุคคลทั่วไป และนิติบุคคล พัฒนาสื่อการเรียนการสอนบนคอมพิวเตอร์แท็บเล็ตหรือเรียกว่า แอปพลิเคชันสำหรับแท็บเล็ต (Application for Tablet) ซึ่งเป็นเครื่องมือที่ สพฐ.พัฒนาขึ้นผ่านเว็บไซต์ www.otpcappcon.com แอปพลิเคชันเพื่อการเรียนรู้เป็นสื่อการศึกษาในรูปแบบดิจิทัลที่เน้นการเรียนรู้ในรูปแบบอินเทอร์เน็ตแอกทีฟ สามารถเรียนรู้ได้ด้วยตนเองและแบบกลุ่มร่วมมือ ช่วยให้ผู้เรียนสามารถสร้างสรรค์ชิ้นงาน ผลงานประกอบการเรียนรู้ หรือสร้างองค์ความรู้ใหม่ และเรียนรู้ผ่านคอมพิวเตอร์แท็บเล็ต [4]

เทคนิคการเรียนรู้แบบเพื่อนคู่คิด (Think- Pair- Share) เป็นวิธีการเรียนแบบร่วมมือในยุคเทคโนโลยีสารสนเทศ สามารถจัดได้ทั้งระบบออนไลน์และออฟไลน์ เพื่อช่วยพัฒนาความรู้ทักษะเพิ่มแรงจูงใจและช่วยแก้ปัญหาต่างๆ ในระหว่างการปฏิบัติได้เป็นอย่างดี สร้างบรรยากาศแห่งการสัมพันธ์และเกิดความภาคภูมิใจ [5]

จากการศึกษากรอบนโยบายเทคโนโลยีสารสนเทศและการสื่อสารเพื่อการศึกษา นโยบายแท็บเล็ตเพื่อการศึกษา แอปพลิเคชันเพื่อการเรียนรู้ การส่งเสริมและพัฒนาครูทำให้ผู้วิจัยได้ส่งเสริมครูพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ต ด้วยเทคนิคเพื่อนคู่คิด โดยมุ่งเน้นให้

ครูมีความรู้ ทักษะและเจตคติ ในการพัฒนาแอปพลิเคชันเพื่อ การเรียนรู้บนคอมพิวเตอร์แท็บเล็ตได้อย่างมีประสิทธิภาพ

2. วัตถุประสงค์

2.1 เพื่อสังเคราะห์องค์ประกอบของการส่งเสริมครูพัฒนา แอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ต ด้วยเทคนิค เพื่อนคู่คิด

2.2 เพื่อสังเคราะห์ขั้นตอนการส่งเสริมครูพัฒนา แอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ต ด้วยเทคนิค เพื่อนคู่คิด

3. วรรณกรรมที่เกี่ยวข้อง

สุภาณี เส็งศรี [6] กล่าวว่า การอบรมครูมี 3 รูปแบบ คือ 1) รูปแบบการอบรมที่เน้นวิทยากรเป็นหลัก 2) รูปแบบ การอบรมที่เน้นสื่อการอบรมและกิจกรรมภาคปฏิบัติ และ 3) รูปแบบการอบรมที่เน้นผู้เข้าอบรมเป็นหลัก โดย กระบวนการอบรมจะมี 3 ขั้นตอน คือ 1) ขึ้นโน้มน้าวและ ไตร่ตรอง 2) ขึ้นฝึกปฏิบัติเพื่อสร้างประสบการณ์ และ 3) ขึ้นนำไปใช้ประโยชน์ สอดคล้องกับ สิริรณาทิ กิจเกื้อกูล [7] กล่าวว่า รูปแบบวิธีการพัฒนาครูตามแนวทางปฏิรูปการศึกษา มี 3 ขั้นตอน คือ 1) ขึ้นเตรียมการ 2) ขึ้นสร้างกิจกรรม การพัฒนาครู และ 3) ขึ้นติดตามสะท้อนผลอย่างต่อเนื่อง

มนต์ชัย เทียนทอง [8] กล่าวว่า Think-Pair-Share ถูก พัฒนาขึ้นโดย Frank Lyman แห่งมหาวิทยาลัยแมริแลนด์ เป็นเทคนิคการเรียนรู้แบบร่วมมือ โดยแบ่งผู้อบรมเป็นคู่ๆ ดำเนินกิจกรรมในลักษณะของคู่ ประกอบด้วย 3 ขั้นตอน คือ 1) Think เป็นการท้าทายให้ผู้อบรมได้คิดและไตร่ตรองจาก คำถามปลายเปิด หรือการเฝ้าสังเกตพฤติกรรม 2) Pair เป็นการ จัดให้ผู้อบรมจับคู่กันเป็นคู่ๆ เพื่อแลกเปลี่ยนความคิดเห็น ซึ่งกันและกันในประเด็นปัญหาที่กำหนดไว้ เพื่อร่วมกันค้นหา ข้อสรุปหรือตอบคำถามที่ต้องการ และ 3) Share เป็นการสลาย จากการจัดกลุ่มกันเป็นคู่ๆ แล้วสรุปผลการค้นหาคำตอบ ร่วมกันเพื่อแลกเปลี่ยนความรู้ สรุปและอภิปรายผลการค้นพบ

การพัฒนาแอปพลิเคชันเพื่อการเรียนรู้เป็นการส่งเสริม การใช้เทคโนโลยีเป็นเครื่องมือยกระดับคุณภาพและกระจาย โอกาสทางการศึกษา ในปี พ.ศ. 2556 สพฐ. ได้จัดอบรมครูและ

ศึกษานิเทศก์ ใน 5 ภูมิภาค จำนวน 1,000 คน ซึ่งแอปพลิเคชัน ที่พัฒนาขึ้นจะแบ่งเป็น 3 ประเภท คือ 1) แอปพลิเคชันรูปแบบ เสริมการเรียนรู้ (Learning Media) เป็นแอปพลิเคชันที่นำเสนอ เนื้อหา เพื่อให้ผู้เรียนสามารถเรียนรู้ได้ด้วยตนเอง 2) แอปพลิเคชันรูปแบบเสริมการสอน (Instruction Media) เป็นแอปพลิเคชันที่ใช้เป็นสื่อช่วยครูในการสอนและ 3) แอปพลิเคชันรูปแบบสร้างองค์ความรู้ (Constructional Media) เป็นแอปพลิเคชันที่ช่วยให้ผู้เรียนสามารถสร้างสรรค์ ชิ้นงาน ผลงานประกอบการเรียนรู้ หรือสร้างองค์ความรู้ใหม่ๆ ผลการวิจัยพบว่าครูมีทัศนคติเชิงบวกต่อการใช้นวัตกรรม เครื่องมือการพัฒนาแอปพลิเคชันที่ สพฐ. พัฒนาขึ้นในระดับ มาก [9] ซึ่งมหาวิทยาลัยราชภัฏมหาสารคาม ได้จัดอบรมครู พัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ต โดย ใช้เครื่องมือ สพฐ. ผ่านเว็บไซต์ www.otpcappcon.com โดย อบรมครูและบุคลากรทางการศึกษาในเขตพื้นที่บริการ จำนวน 181 คน ผลการอบรมพบว่า ผู้เข้ารับการอบรมสามารถพัฒนา แอปพลิเคชันที่มีคุณภาพ จำนวน 172 คน คิดเป็นร้อยละ 95.03 ของผู้เข้าอบรมและผู้เข้าอบรมมีความพึงพอใจในระดับ มากที่สุด [10]

4. วิธีการดำเนินการวิจัย

4.1 กลุ่มเป้าหมาย ที่ใช้ในการวิจัยครั้งนี้เป็นผู้เชี่ยวชาญ ที่สำเร็จการศึกษาระดับปริญญาเอก และมีประสบการณ์ใน การพัฒนาแอปพลิเคชัน จากเขตพื้นที่การศึกษาและมหาวิทยาลัย จำนวน 12 คน โดยแบ่งเป็น 4 กลุ่มๆ ละ 3 คน คือ 1) กลุ่ม นักเทคโนโลยีสารสนเทศ 2) กลุ่มผู้บริหาร โรงเรียนและ ศึกษานิเทศก์ 3) กลุ่มวิทยากรด้านการพัฒนาแอปพลิเคชัน และ 4) กลุ่มนักวิชาการศึกษามากจากมหาวิทยาลัย

4.2 เครื่องมือที่ใช้ในการวิจัย

แบบสอบถามความคิดเห็นของผู้เชี่ยวชาญที่มีต่อ องค์ประกอบและขั้นตอนของการส่งเสริมครูพัฒนาแอปพลิเคชัน เพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ต ด้วยเทคนิคเพื่อนคู่คิด

4.3 ขั้นตอนการวิจัย

4.3.1 ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับการส่งเสริม พัฒนาครู การเรียนรู้แบบเพื่อนคู่คิด การพัฒนาแอปพลิเคชัน คอมพิวเตอร์แท็บเล็ต

4.3.2 สํารวจความต้องการในการพัฒนาส่งเสริมการเรียนรู้ของครู จากครูสังกัดสำนักงานการศึกษามัธยมศึกษา เขต 24 กาฬสินธุ์

4.3.3 สรุปข้อมูลที่ได้จากการศึกษาเอกสาร และจากการสำรวจความต้องการของครู นำเสนออาจารย์ที่ปรึกษา

4.3.4 สั้งเคราะห์องค์ประกอบของการส่งเสริมครูพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ต ด้วยเทคนิคเพื่อนคู่คิด

4.3.5 สั้งเคราะห์ขั้นตอนการส่งเสริมครูพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ต ด้วยเทคนิคเพื่อนคู่คิด

4.3.6 นำเสนอร่างองค์ประกอบ และร่างขั้นตอนการส่งเสริมครู ต่ออาจารย์ที่ปรึกษาเพื่อปรับปรุงแก้ไข

4.3.7 ประชุมกลุ่มย่อยผู้เชี่ยวชาญ จำนวน 12 คน เพื่อสอบถามความคิดเห็นที่มีต่อร่างองค์ประกอบ และร่างขั้นตอนการส่งเสริมครู

4.3.8 วิเคราะห์และสรุปผลการวิจัย

4.4 การวิเคราะห์ข้อมูล

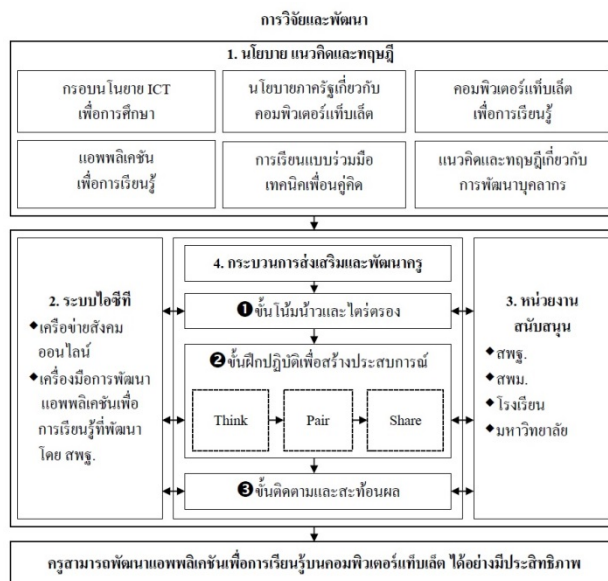
การวิเคราะห์ข้อมูล ผู้วิจัยนำแบบสอบถามที่ได้จากกลุ่มเป้าหมายมาวิเคราะห์ระดับความคิดเห็นโดยใช้ ค่าเฉลี่ย (\bar{X}) และส่วนเบี่ยงเบนมาตรฐาน (S.D.) โดยนำค่าเฉลี่ยที่ได้เทียบกับเกณฑ์การประเมินดังนี้ [11]

- 4.51 – 5.00 หมายความว่า เหมาะสมมากที่สุด
- 3.51 – 4.50 หมายความว่า เหมาะสมมาก
- 2.51 – 3.50 หมายความว่า เหมาะสมปานกลาง
- 1.51 – 2.50 หมายความว่า เหมาะสมน้อย
- 1.00 – 1.50 หมายความว่า เหมาะสมน้อยที่สุด

5. ผลการดำเนินงาน

5.1 ผลการสังเคราะห์องค์ประกอบของการส่งเสริมครูพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ต ด้วยเทคนิคเพื่อนคู่คิด

5.1.1 องค์ประกอบของการส่งเสริมครูพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ต ด้วยเทคนิคเพื่อนคู่คิด ประกอบด้วยองค์ประกอบ จำนวน 4 ส่วน ดังภาพที่ 1



ภาพที่ 1 : องค์ประกอบของการส่งเสริมครูพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ต ด้วยเทคนิคเพื่อนคู่คิด

จากภาพที่ 1 องค์ประกอบแต่ละส่วนมีรายละเอียด ดังนี้

1) นโยบายแนวคิดทฤษฎี ประกอบด้วย กรอบนโยบาย ICT เพื่อการศึกษา นโยบายภาครัฐเกี่ยวกับคอมพิวเตอร์แท็บเล็ต คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ แอปพลิเคชันเพื่อการเรียนรู้ แนวคิดทฤษฎีเกี่ยวกับการพัฒนาบุคลากร และการเรียนรู้แบบร่วมมือเทคนิคเพื่อนคู่คิด

2) ระบบเทคโนโลยีสารสนเทศและการสื่อสาร (ไอซีที) ประกอบด้วย เครื่องข่ายสังคมออนไลน์ และเครื่องมือการพัฒนาแอปพลิเคชันเพื่อการเรียนรู้ที่พัฒนาโดย สฟฐ.

3) หน่วยงานสนับสนุน มี 4 หน่วยงาน คือ สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน สำนักงานเขตพื้นที่การศึกษามัธยมศึกษา โรงเรียน และมหาวิทยาลัย

4) กระบวนการส่งเสริมและพัฒนาครูมี 3 ขั้นตอน คือ (1) ขึ้นโน้มน้าวและไต่รตรง (2) ขึ้นฝึกปฏิบัติเพื่อสร้างประสบการณ์ จะใช้เทคนิคเพื่อนคู่คิด (Think-Pair-Share) และ (3) ขึ้นติดตามและสะท้อนผล

5.1.2 ผลการศึกษาความคิดเห็นของผู้เชี่ยวชาญที่มีต่อองค์ประกอบของการส่งเสริมครู ผู้วิจัยนำขั้นตอนและวิธีการส่งเสริมครูเข้าสู่การประชุมกลุ่มย่อยเพื่อให้ผู้เชี่ยวชาญร่วมพิจารณาความเหมาะสมของขั้นตอนและวิธีการ และประเมินขั้นตอนและวิธีการส่งเสริมครู แสดงในตารางที่ 1

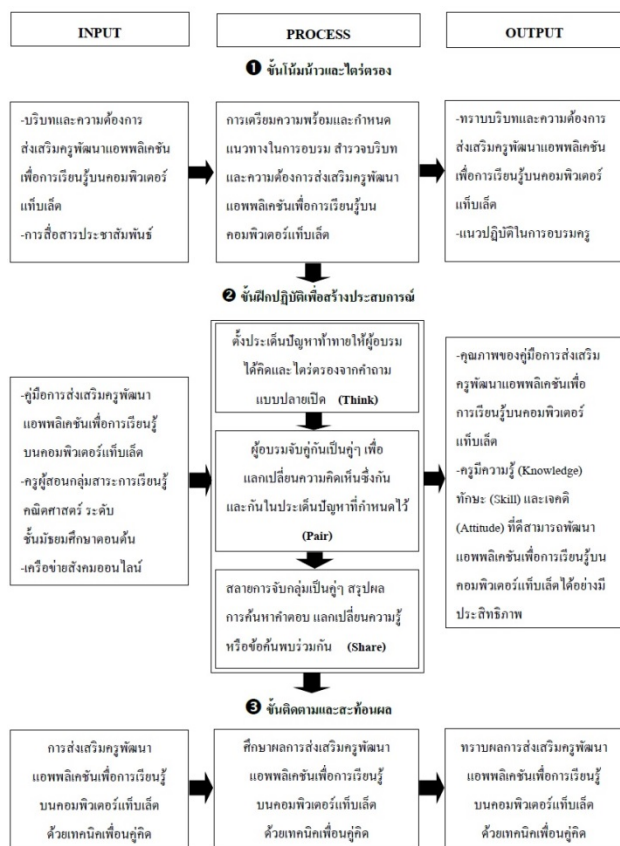
ตารางที่ 1: ผลการศึกษาความคิดเห็นของผู้เชี่ยวชาญที่มีต่อองค์ประกอบของการส่งเสริมครู

รายการ	ระดับความคิดเห็น		
	\bar{X}	S.D.	ความหมาย
1. นโยบาย แนวคิด และทฤษฎี	4.76	0.44	มากที่สุด
2. ระบบเทคโนโลยีสารสนเทศและการสื่อสาร	4.79	0.39	มากที่สุด
3. หน่วยงานสนับสนุน	4.87	0.29	มากที่สุด
4. กระบวนการส่งเสริมและพัฒนาครู	4.83	0.57	มากที่สุด
รวมเฉลี่ย	4.75	0.43	มากที่สุด

จากตารางที่ 1 พบว่า ภาพรวมของความคิดเห็นของผู้เชี่ยวชาญที่มีต่อองค์ประกอบ มีความเหมาะสมอยู่ในระดับมากที่สุด ($\bar{X} = 4.75$, S.D. = 0.43) เมื่อพิจารณาเป็นรายด้านพบว่าทุกด้านมีความเหมาะสมในระดับมากที่สุด

5.2 ผลการสังเคราะห์ขั้นตอนการส่งเสริมครูพัฒนา แอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ต ด้วยเทคนิคเพื่อนคู่คิด

5.2.1 ขั้นตอนการส่งเสริมครูพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ต ด้วยเทคนิคเพื่อนคู่คิด ประกอบด้วย 3 ขั้นตอน ดังแสดงในภาพที่ 2



ภาพที่ 2: ขั้นตอนการส่งเสริมครูพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ต ด้วยเทคนิคเพื่อนคู่คิด

จากภาพที่ 2 รายละเอียดกระบวนการส่งเสริมและพัฒนาครู มี 3 ขั้นตอน ดังนี้

1) ขั้นโน้มน้าวและไตร่ตรอง เป็นขั้นเตรียมความพร้อมและกำหนดแนวทางในการอบรม เพื่อให้ทราบบริบทและ ความต้องการส่งเสริมครูพัฒนาแอปพลิเคชันเพื่อ การเรียนรู้ บนคอมพิวเตอร์แท็บเล็ต

2) ขั้นฝึกปฏิบัติเพื่อสร้างประสบการณ์ เป็นขั้นตอน จัดประสบการณ์เพื่อการวิเคราะห์ออกแบบและพัฒนา แอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ตซึ่งกระบวนการส่งเสริมและพัฒนาครูจะใช้เทคนิคเพื่อนคู่คิด (Think-Pair-Share)

3) ขั้นติดตามและสะท้อนผล เป็นขั้นตอนการติดตามที่มุ่งผลสำเร็จ ครูมีการเปลี่ยนแปลงพฤติกรรมและมั่นใจต่อการนำความรู้ ทักษะ ที่ได้จากการพัฒนาไปปฏิบัติในสภาพจริง

5.2.2 ผลการศึกษา ความคิดเห็นของผู้เชี่ยวชาญที่มีต่อขั้นตอนการส่งเสริมครูพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ต ด้วยเทคนิคเพื่อนคู่คิด ผู้วิจัยนำขั้นตอนการส่งเสริมครูเข้าสู่การประชุมกลุ่มย่อยเพื่อให้ผู้เชี่ยวชาญร่วมพิจารณาความเหมาะสมของขั้นตอนและประเมินขั้นตอน การส่งเสริมครูพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ต ด้วยเทคนิคเพื่อนคู่คิด ดังแสดงใน ตารางที่ 2

ตารางที่ 2: ความคิดเห็นของผู้เชี่ยวชาญที่มีต่อขั้นตอนการส่งเสริมครูพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ต ด้วยเทคนิคเพื่อนคู่คิด

รายการ	ระดับความคิดเห็น		
	\bar{X}	S.D.	ความหมาย
1. ขั้นโน้มน้าวและไตร่ตรอง	4.75	0.45	มากที่สุด
2. ขั้นฝึกปฏิบัติเพื่อสร้างประสบการณ์	4.77	0.38	มากที่สุด
3. ขั้นติดตามและสะท้อนผล	4.75	0.39	มากที่สุด
รวมเฉลี่ย	4.75	0.40	มากที่สุด

จากตารางที่ 2 พบว่า ความคิดเห็นของผู้เชี่ยวชาญที่มีต่อขั้นตอนการส่งเสริมครูพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ต ด้วยเทคนิคเพื่อนคู่คิด ภาพรวมอยู่ในระดับมากที่สุด ($\bar{X} = 4.75$, S.D. = 0.40) เมื่อพิจารณารายด้านพบว่าทุกด้านมีความเหมาะสมอยู่ในระดับมากที่สุด

6. สรุป

การวิจัยในครั้งนี้ องค์ประกอบของการส่งเสริมการพัฒนา แอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ต ด้วยเทคนิคเพื่อนคู่คิด มีองค์ประกอบที่สำคัญ 4 ส่วน คือ 1) นโยบาย แนวคิดและทฤษฎี 2) ระบบเทคโนโลยีสารสนเทศและการสื่อสาร 3) หน่วยงานสนับสนุน และ 3) กระบวนการส่งเสริมและพัฒนาครู โดยที่ขั้นตอนการส่งเสริมครูพัฒนา แอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ต ด้วยเทคนิคเพื่อนคู่คิด มี 3 ขั้นตอน คือ 1) ขึ้นโน้มน้าวและได้ตรงต่อ 2) ขึ้นฝึกปฏิบัติเพื่อสร้างประสบการณ์ จะใช้เทคนิคเพื่อนคู่คิด (Think-Pair-Share) และ 3) ขึ้นติดตามและสะท้อนผล ทั้งนี้ผู้วิจัยจะนำองค์ประกอบและขั้นตอนการส่งเสริมครูพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ต ด้วยเทคนิคเพื่อนคู่คิด ที่สังเคราะห์ไปสู่การดำเนินการวิจัยต่อไป

เอกสารอ้างอิง

- [1] กระทรวงเทคโนโลยีสารสนเทศและการสื่อสาร, “กรอบนโยบายเทคโนโลยีสารสนเทศและการสื่อสารระยะ พ.ศ. 2554-2563 ของประเทศไทย,” [ออนไลน์]. <http://www.ict2020.in.th/?q=content/national-ict-policy-framework-2011-2020-ict-2020> [1 เมษายน 2555].
- [2] สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน, “คู่มืออบรมสร้างสื่อการเรียนรู้สู่แท็บเล็ต,” กรุงเทพฯ : สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน, 2556.
- [3] Xiang , W. and Colleagues, “ Use of Wireless Tablets PCs as an Effective Learning and Teaching Enhancement Tool,” [online]. Available from: <http://www.caudit.edu.au/educauseaustralia09/pdf> [2014,January 16].
- [4] El-Gayar , O. ; Moran , M. and Hawkes , M, “Student’s Acceptance of Tablet PC and Implications for Education Institutions,” [online]. Available from: <http://www.ifets.info/journals/14.2/5.pdf> [2014,January 16].
- [5] ส่องศรี โปธิ์ทอง, “การศึกษาผลของการพัฒนาบุคลากรด้วยการจัดการความรู้แบบเพื่อนช่วยเพื่อนต่อผลสัมฤทธิ์ด้านการใช้เทคโนโลยีสารสนเทศและการสื่อสารเพื่อปฏิบัติราชการแนวใหม่ของสำนักผู้ตรวจราชการประจำเขตตรวจราชการที่ 11,” *วารสารบัณฑิตศึกษา มหาวิทยาลัยราชภัฏสงขลา* ฉบับที่ 1 ปีที่ 1 มกราคม-มิถุนายน 2551 หน้า 133-141.
- [6] สุภาณี เส็งศรี, “รูปแบบกระบวนการฝึกอบรมครูประจำการเพื่อพัฒนาสื่ออิเล็กทรอนิกส์,” *วารสารศึกษาศาสตร์ มหาวิทยาลัยนครสวรรค์* ปีที่ 13 ฉบับที่ 3 กันยายน-ธันวาคม 2554 หน้า 173-188.
- [7] สิริรณภ กิจเกื้อกุล, “การสังเคราะห์รูปแบบการพัฒนาครูตามแนวปฏิรูป : ประสบการณ์จากวิทยานิพนธ์ พุทธศักราช 2543-2551,” *วารสารศึกษาศาสตร์ มหาวิทยาลัยนครสวรรค์* ปีที่ 12 ฉบับที่ 2 พฤษภาคม-สิงหาคม 2553 หน้า 45-59.
- [8] มนต์ชัย เทียนทอง, “เทคนิคการเรียนรู้แบบร่วมมือแบบ Mentor Coached Think-Pair-Share เพื่อเพิ่มผลสัมฤทธิ์ทางการเรียนในการเรียนรู้ออนไลน์,” *วารสารวิชาการพระจอมเกล้าพระนครเหนือ* ปีที่ 18 ฉบับที่ 1 มกราคม-เมษายน 2551 หน้า 99-105.
- [9] สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน, “การศึกษาผลการอบรมครูโดยใช้นวัตกรรมระบบพัฒนาสื่อการเรียนการสอนตามโครงการประกวดสร้างสื่อการเรียนรู้สู่แท็บเล็ต,” กรุงเทพฯ : สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน, 2556.
- [10] มหาวิทยาลัยมหาสารคาม, “รายงานการอบรมโครงการศูนย์ทางไกลฯ มหาวิทยาลัยมหาสารคาม,” *มหาสารคาม : คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏมหาสารคาม*, 2556.
- [11] บุญชม ศรีสะอาด, “การวิจัยเบื้องต้น,” พิมพ์ครั้งที่ 7, กรุงเทพฯ : สุวีริยาสาส์น, 2545.

แนวทางการพัฒนารูปแบบการจัดการเรียนการสอนรายวิชาออปแอมป์และลิเนียร์ไอซีด้วยเทคโนโลยีสื่อสารทางไกลและเทคโนโลยีสารสนเทศ

Guidelines for Developing Teaching & Learning Model of Op-Amp and Linear IC by using Distance Communications Technology and Information Technology

ปิติภักย์ ปิ่นรอด (Pitipark Pinrod)¹ และ ประชยนันท์ นิลสุข (Prachyanun Nilsook)²

¹แผนกวิชาอิเล็กทรอนิกส์ คณะไฟฟ้าอิเล็กทรอนิกส์ สถาบันการอาชีวศึกษาภาคเหนือ1 วิทยาลัยเทคนิคลำพูน

²ภาควิชาครุศาสตร์เทคโนโลยี คณะครุศาสตร์อุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹funnyfm@gmail.com, ²prachyanun@hotmail.com

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนารูปแบบการจัดการเรียนการสอนและศึกษาประสิทธิภาพรูปแบบการจัดการเรียนการสอนรายวิชาออปแอมป์และลิเนียร์ไอซีด้วยเทคโนโลยีสื่อสารทางไกลและเทคโนโลยีสารสนเทศ ประชากรในการวิจัยครั้งนี้คือ ผู้ทรงคุณวุฒิด้านเทคโนโลยีการศึกษา โดยการเลือกแบบเจาะจงจำนวน 5 คน วิเคราะห์ข้อมูลโดยใช้ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ผลการวิจัย พบว่า รูปแบบการเรียนการสอนรายวิชาออปแอมป์และลิเนียร์ไอซีด้วยเทคโนโลยีสื่อสารทางไกลและเทคโนโลยีสารสนเทศ HCL Model (Hyflex Constructionism Model) เป็นรูปแบบการจัดการเรียนการสอนแบบยืดหยุ่นตามสภาพแวดล้อมการเรียนรู้ใช้กระบวนการจัดการเรียนการสอนที่ให้ผู้เรียนสร้างองค์ความรู้ด้วยตนเอง ซึ่งผู้เรียนสามารถเรียนได้ทั้งในห้องเรียน นอกห้องเรียน ทุกที่ ทุกเวลา ประกอบด้วย องค์ประกอบหลัก 4 ประการ คือ 1.ปัจจัยนำเข้า(INPUT) 2.ปัจจัยด้านกระบวนการ(PROCESS) 3.ปัจจัยนำออก(OUTPUT) และ 4.ปัจจัยย้อนกลับ(FEEDBACK) ประกอบด้วยขั้นตอนการจัดการเรียนการสอน 11 ขั้นตอนหลัก และ 23 ขั้นตอนย่อย คือ 1.ขั้นการวิเคราะห์(Analysis) 1.1) วิเคราะห์หลักสูตร 1.2) วิเคราะห์เนื้อหารายวิชา 1.3) วิเคราะห์ผู้เรียน 1.4) วิเคราะห์ระบบพื้นฐาน 2.ขั้นการออกแบบและพัฒนาระบบ (Design & Development) 2.1) กำหนดเป้าหมายการเรียนรู้ 2.2) พัฒนาวัตถุประสงค์การสอน 2.3) รวบรวมและ

ออกแบบสร้างสรรค์เนื้อหา 2.4) เลือกกิจกรรมการสอนให้ผู้เรียนมีส่วนร่วมในการเรียนรู้ 2.5) สร้างสรรค์วิธีการสื่อสารเพื่อการสอน 2.6) เลือกวิธีการถ่ายทอดเนื้อหา 2.7) เตรียมแหล่งสนับสนุนการเรียนรู้และแหล่งข้อมูลอ้างอิงที่เกี่ยวข้อง 3.ขั้นการเตรียมการ(Preparation) 3.1) การปฐมนิเทศ 3.2) รับผิดชอบต่อเป้าหมายหรือวัตถุประสงค์ 3.3) ขอบข่าย 4.ขั้นแสวงหาความรู้ (Searching) 5.ขั้นรวบรวมและจัดการกับข้อมูลที่เกี่ยวข้อง (Data collection) 6.ขั้นวางแผนสร้างชิ้นงาน(Planning) 7.ขั้นปฏิบัติการสร้างชิ้นงาน (Implementation) 7.1) การสืบค้น 7.2) การทดลองทำ 7.3) ผลจากการทดลอง 7.4) การตรวจสอบย้อนกลับดู 8.ขั้นการประเมินผล(Evaluation) 8.1) ด้านความคืบหน้า(Progress) 8.2) ด้านกระบวนการ (Process) 8.3) ด้านผลงาน (Product) 9.ขั้นนำเสนอผลงาน (Presentation) 10) ชิ้นเผยแพร่ผลงาน (Published Works) 11) ขั้นการตรวจสอบ (Check) 11.1) ด้านความคืบหน้า(Progress) และ 11.2) ด้านกระบวนการ (Process) ผลการหาประสิทธิภาพของรูปแบบจากผู้ทรงคุณวุฒิด้านเทคโนโลยีการศึกษา พบว่า รูปแบบที่พัฒนาขึ้นมีคุณภาพในระดับมากที่สุด ที่ค่าเฉลี่ย 4.75 ค่าเบี่ยงเบนมาตรฐาน S.D. 0.07

คำสำคัญ: เทคโนโลยีสื่อสารทางไกล เทคโนโลยีสารสนเทศ รูปแบบการจัดการเรียนการสอนออปแอมป์และลิเนียร์ไอซี

Abstract

This research were aim to developed teaching & leaning model and studied effective of teaching & learning model of Op-Amp and Linear IC by using distance communications technology and information technology. Population in this research was 5 experts in field of education technology by purposive sampling. Data analysis were used statistic by mean and standard deviation. The results of this research were found as the teaching and learning model: HCL model (Hyflex Constructionism Model) is a flexibility model for learning environment and emphasized on the learners construct knowledge for themselves, which can learn in both of in and outside the classroom, anywhere and anytime. The model consists with 4 components as 1.Inputs 2.process 3. Output 4.feedback with 11 mainly steps and 23 sub-steps as 1) Analysis 1.1) curriculum analysis 1.2) content analysis 1.3) learners analysis 1.4) basis system analysis 2) design and development 2.1) learning targeted 2.2) objective development 2.3) collecting, designing and creative content. 2.4) identify learning activities as learner participated 2.5) created teaching & learning communication methods 2.6) identify method for content transfer. 2.7) identify education resource and related reference resources 3) preparation 3.1) orientation. 3.2) acknowledge the target and objective 3.3) Scope 4). knowledge searching 5) data collection 6) Planning 7) Implementation 7.1) search 7.2) experimental 7.3) experiment result 7.4) traceability 8. evaluation 8.1) in progress 8.2) process 8.3) product) 9) presented 10) published works 11) Check 11.1) progress and 11.2) process. The results of the performance of the model from the experts in field of education technology was showed as the teaching & learning on Op-Amp and Linear IC by using distance communications technology and information technology which created in this research has the highest quality at 4.75of mean (\bar{x}) and 0.07 of standard deviation ($SD = 0.07$).

Key words: Distance Education, Information Technology, Teaching and learning model of op amp and

linear integrated Circuits with distance communications technology and information technology

1. บทนำ

สำนักงานคณะกรรมการการอาชีวศึกษาดำเนินการจัดการศึกษาที่เน้นให้มีการปฏิรูปการศึกษาทั้งระบบ โดยมีการปฏิรูปการศึกษาในทศวรรษที่สองให้คนไทยได้เรียนรู้ตลอดชีวิตอย่างมีคุณภาพตามเป้าหมาย 3 ประการคือ พัฒนาคุณภาพมาตรฐานการศึกษาและการเรียนรู้ของคนไทย เพิ่มโอกาสทางการศึกษาและเรียนรู้อย่างทั่วถึงและมีคุณภาพ และส่งเสริมการมีส่วนร่วมของทุกภาคส่วนของสังคมภายใต้กรอบแนวทางการปฏิรูปการศึกษา 4 ประการ คือ ด้านพัฒนาคุณภาพคนไทยยุคใหม่ พัฒนาคุณภาพครูยุคใหม่ พัฒนาคุณภาพสถานศึกษาและแหล่งเรียนรู้ยุคใหม่ และพัฒนาคุณภาพการบริหารจัดการใหม่ (สำนักงานคณะกรรมการการอาชีวศึกษา, 2555) ^[1] งานวิจัยอนาคตภาพการอาชีวศึกษาไทยในทศวรรษหน้า(พ.ศ.2554 - พ.ศ.2564) พบว่าการจัดการเรียนการสอนหลักสูตรอาชีวศึกษาต้องตอบสนองความต้องการของแรงงานได้หลากหลายทั้งในและนอกระบบ ระบบทวิภาคีเทียบโอนประสบการณ์และระบบการศึกษาทางไกล วิธีการเรียนการสอนยืดหยุ่นตามเทคโนโลยีความต้องการของผู้เรียนและสถานประกอบการ (ดวงนภามกรานุรักษ์, 2554)^[2]

การจัดการเรียนการสอนด้วยเทคโนโลยีศตวรรษที่ 21 สามารถนำไปประยุกต์กับการจัดการศึกษาระบบทวิภาคี และการจัดการศึกษาแบบเทียบโอนประสบการณ์สำหรับผู้เรียนอาชีวศึกษา โดยจัดการเรียนการสอนที่มีความเหมาะสมกับการเปลี่ยนแปลงของเทคโนโลยีในปัจจุบัน ซึ่งผู้วิจัยมุ่งหวังที่จะพัฒนารูปแบบการจัดการเรียนการสอนด้วยเทคโนโลยีสื่อสารทางไกลและเทคโนโลยีสารสนเทศให้ผู้เรียนสามารถเรียนรู้ได้ทุกที่ทุกเวลาสู่การเรียนรู้ตลอดชีวิต (Long life Learning) การจัดการเรียนการสอนด้วยเทคโนโลยีสื่อสารทางไกลและเทคโนโลยีสารสนเทศ จึงเป็นทางเลือกใหม่ในการนำเทคโนโลยีมาใช้เพื่อพัฒนาและเพิ่มประสิทธิภาพการศึกษา ผู้วิจัยจึงได้นำแนวคิดการจัดการเรียนการสอนตามแนวทางการจัดการศึกษาข้างต้นมาใช้ ในการจัดการเรียนการสอนในวิชาออปแอมป์และลิเนียร์ไอซี ด้วยเทคโนโลยีสื่อสารทางไกลและเทคโนโลยีสารสนเทศ ให้สอดคล้องกับการปฏิรูปการศึกษาใน

ทศวรรษที่สอง (พ.ศ. ๒๕๕๒-๒๕๖) ให้ผู้เรียนได้เรียนรู้ตลอดชีวิต ทุกที่ทุกเวลา

2. วรรณกรรมที่เกี่ยวข้อง

2.1. การออกแบบและพัฒนารูปแบบ ADDIE MODEL

ADDIE MODEL เป็นการออกแบบระบบการเรียนการสอน โดยพัฒนาโปรแกรมการสอนจากจุดเริ่มต้นจนถึงสุด (Barbara Seel, Zita Glasgow (1998) อ้างใน ฉัตรพงศ์ ชูแสงนิล, 2555)^[3] มี 5 ขั้นตอน คือ 1.การวิเคราะห์ 2.การออกแบบ 2.1)การออกแบบบทเรียน 2.2) การออกแบบผังงานและบทดำเนินเรื่อง 2.3) การออกแบบหน้าจอภาพ 3. การพัฒนาเพื่อสร้าง/เขียนโปรแกรมและผลิตเอกสารประกอบการเรียน 3.1)การเตรียมองค์ประกอบข้อความ, ภาพ, เสียง และโปรแกรมจัดการบทเรียน 3.2) สร้างบทเรียนด้วยโปรแกรมคอมพิวเตอร์เปลี่ยน story board เป็นบทเรียนคอมพิวเตอร์ช่วยสอน 3.3) สร้างเอกสารประกอบการเรียน 4.การนำไปใช้ 5.การประเมินผล 5.1) การประเมินผลเพื่อพัฒนา 5.2)การประเมินผลรวม จากนั้นนำสื่อต้นแบบ (Prototype) ไปประเมินผลเพื่อพัฒนาโดยให้ผู้เชี่ยวชาญประเมินสื่อต้นแบบก่อนนำไปทดลองกับกลุ่มทดลองและนำเข้าสู่ขั้นตอนการทดลองใช้กับผู้เรียน

2.2 การจัดการเรียนการสอนแบบยืดหยุ่น (HyFlex Learning)

การเรียนการสอนแบบยืดหยุ่น (HyFlex Learning) คือ การเรียนรู้ผสมผสานแบบยืดหยุ่น เป็นเทคโนโลยีการศึกษาที่ช่วยเสริมสร้างประสบการณ์ในการจัดการเรียนการสอนระหว่างผู้สอนกับผู้เรียนคือรูปแบบการเรียนการสอนที่ผสมผสานระหว่างการเรียนการสอนแบบเผชิญหน้ากับการเรียนออนไลน์ โดยผู้เรียนมีส่วนร่วมกิจกรรมในชั้นเรียน ห้องเรียนออนไลน์หรือทั้งสองกิจกรรม เป็นการเพิ่มโอกาสทางการศึกษา และส่งเสริมการมีส่วนร่วมของผู้เรียน (Beatty, 2012)^[4] สอดคล้องกับแนวคิดด้านนวัตกรรมเทคโนโลยีการศึกษาทางไกลที่มุ่งเน้นประสิทธิภาพการจัดการศึกษา ก่อเกิดการเรียนรู้และประหยัดทรัพยากร สามารถเรียนรู้ได้ด้วยตนเองผ่านเทคโนโลยีคอมพิวเตอร์และอินเทอร์เน็ต เป็นนวัตกรรมและเทคโนโลยีที่ผสมผสาน และสามารถตอบสนองผู้เรียนเป็นสำคัญ การปฏิรูปการศึกษาในทศวรรษที่สอง(พ.ศ. 2552-2561)ที่ต้องการให้คนไทยได้เรียนรู้ตลอดชีวิตอย่างมีคุณภาพ สรุปหลักการได้ดังนี้ 1. มุ่งเน้นผู้เรียนเป็นสำคัญเป็นแนวคิดแสวงหาความรู้ด้วยการนำ

ตนเองและรับผิดชอบต่อตนเอง 2.การออกแบบสารเพื่อสื่อความหมายและการเรียนรู้ที่มีประสิทธิภาพ 3.ใช้เวลาเพื่อการเรียนรู้สามารถเรียนรู้ได้ทุกที่ทุกเวลา 4.การขยายโอกาสทางการศึกษาและการเรียนรู้เป็นการเพิ่มโอกาสทางการศึกษาและการเรียนรู้ให้ทั่วถึงและมีคุณภาพ 5.แนวคิดด้านความพร้อม มีการจัดลำดับขั้นของกิจกรรมให้หลากหลายเหมาะสมกับความพร้อมของผู้เรียน ดังนี้ 1.ผู้เรียนมีทางเลือก เลือกวิธีการเรียนของตนเองเพื่อให้บรรลุเป้าหมายของรายวิชาที่กำหนดไว้ 2.ความเท่าเทียมกัน มีโอกาสได้เรียนรู้โดยความเสมอภาคกัน 3. ความสามารถในการนำกลับมาใช้ใหม่ การใช้สื่อการเรียนการสอนเพื่อถ่ายทอดความรู้ทั้งแบบออนไลน์และแบบเผชิญหน้าหรืออื่นๆ เป็นวัสดุที่นำกลับมาใช้ซ้ำใหม่ได้กับผู้เรียนทุกคน 4. ความสามารถในการเข้าถึงบทเรียนได้ง่ายไม่ยุ่งยากซับซ้อนด้วยศักยภาพ ของเทคโนโลยีอินเทอร์เน็ต

2.3 การจัดการเรียนการสอนตามแนว Constructionism

การจัดการเรียนการสอนตามแนว Constructionism (ทิสนา แจมมณี, 2551)^[5] เน้นผู้เรียนเป็นฝ่ายสร้างความรู้ขึ้นด้วยตนเองทำให้ผู้เรียนต้องใช้ความคิดกระตือรือร้นและความรับผิดชอบต่อการเรียนรู้ของตนเอง ด้วยหลัก 4 ประการ คือ 1. ผู้เรียนสร้างองค์ความรู้ด้วยตนเอง โดยให้ผู้เรียนประกอบกิจกรรมการเรียนรู้ เข้าใจในตนเอง มองเห็นความสำคัญในสิ่งที่เรียนรู้และสามารถเชื่อมโยงความรู้สร้างเป็นองค์ความรู้ใหม่ขึ้นมา 2.ครูจัดบรรยากาศการเรียนการสอนที่เปิดโอกาสให้ผู้เรียนปฏิบัติกิจกรรมการเรียนรู้ด้วยตนเองโดยมีทางเลือกที่หลากหลาย เชื่อมโยงความรู้ใหม่กับความรู้เก่าได้ โดยมีครูช่วยเหลือและอำนวยความสะดวก 3.เรียนรู้จากประสบการณ์และสิ่งแวดล้อมที่เรียนรู้ร่วมกัน 4.ใช้เทคโนโลยีเป็นเครื่องมือแสวงหาคำตอบจากแหล่งความรู้ต่างๆด้วยตนเองเป็นผลให้เกิดพฤติกรรม "เรียนรู้ว่าจะเรียนรู้ได้อย่างไร

2.4 งานวิจัยที่เกี่ยวข้อง

(อุทิศ บำรุงชีพ, 2551)^[6] ได้วิจัยรูปแบบการเรียนการสอนผ่านเว็บตามแนวตามแนวคอนสตรัคชันนิซึมเพื่อพัฒนาความคิดสร้างสรรค์พบว่ารูปแบบการเรียนการสอน ประกอบด้วย 9 ขั้นตอน ได้แก่ (1) ขั้นเตรียมการ (2) ขั้นสืบเสาะค้นคว้า (3) ขั้นรวบรวมข้อมูล(4) ขั้นระดมสมอง ออกิปรายเพื่อเลือกหัวข้อและวางแผน (5) ขั้นลงมือปฏิบัติ (6) ขั้นพิจารณาไตร่ตรองและ

ปรับปรุง (7) ชั้นประเมินผล (8) ชั้นนำเสนอ (9) ชั้นเผยแพร่ผล
แห่งความสำเร็จ และรูปแบบการเรียนการสอนที่สร้างขึ้นมี
ประสิทธิภาพ 84.55/85.50

3. วิจัยดำเนินการวิจัย

แนวคิดและรูปแบบวิธีการดำเนินการวิจัย

ในการศึกษาค้นคว้าครั้งนี้ผู้ศึกษาได้ออกแบบระบบการเรียนการสอนตามแนวทางของ ADDIE MODEL และการใช้เทคโนโลยีสื่อสารทางไกลและเทคโนโลยีสารสนเทศมาจัดการเรียนการสอนแบบยืดหยุ่น (HyFlex Learning) และประยุกต์ใช้หลักการเรียนการสอนในแบบของ Constructionism โดยเน้นให้ผู้เรียนแสวงหาความรู้และรับผิดชอบตนเอง ผสมผสาน สื่อการเรียนการสอน วิธีการและเทคนิคการสอนแบบผ่านชั้นเรียนออนไลน์ด้วยระบบการสื่อสารทางไกลและระบบสารสนเทศ จัดการเรียนรู้ที่นักเรียนและผู้สนใจสามารถเชื่อมโยงการเรียนรู้และประสบการณ์สู่การเรียนรู้ทั้งในฐานะผู้เรียนและผู้สอนได้อย่างต่อเนื่องตลอดเวลาและตลอดชีวิต โดยมีขั้นตอนการวิจัยดังนี้

1. ศึกษา รูปแบบการจัดการเรียนการสอนด้วยเทคโนโลยีสื่อสารทางไกลและเทคโนโลยีสารสนเทศ 1.1) ศึกษา วิเคราะห์ สังเคราะห์ เอกสาร และงานวิจัยที่เกี่ยวข้องด้านสื่อและเทคโนโลยีการศึกษา ทั้งข้อมูลปฐมภูมิ ข้อมูลทุติยภูมิ 1.2) เข้าร่วมประชุมวิชาการ ประชุมเชิงปฏิบัติการ การร่วมอภิปราย การสัมภาษณ์ผู้เชี่ยวชาญ การสังเกตอย่างมีส่วนร่วม การสนทนากลุ่ม และการสัมภาษณ์เชิงลึกด้านสื่อและเทคโนโลยีการศึกษาทั้งระดับชาติและระดับนานาชาติ เพื่อเปรียบเทียบวิธีดำเนินงานที่ทำแล้วในต่างประเทศ และภายในประเทศ ศึกษาเอกสารงานวิจัยที่เกี่ยวข้อง ใช้แบบสังเคราะห์งานวิจัยที่ผ่านการตรวจสอบเครื่องมือจากที่ปรึกษา โครงการวิจัย และใช้สถิติการวิเคราะห์เนื้อหาให้ได้ร่างรูปแบบการจัดการเรียนการสอนด้วยเทคโนโลยีสื่อสารทางไกลและเทคโนโลยีสารสนเทศ

2. ระบุเงื่อนไขและพัฒนา รูปแบบการจัดการเรียนการสอน รายวิชาออนไลน์และลิเนียร์ไอซีด้วยเทคโนโลยีสื่อสารทางไกลและเทคโนโลยีสารสนเทศ 2.1) ระบุเงื่อนไขและพัฒนา รูปแบบ โดยนำร่างรูปแบบการจัดการเรียนการสอนรายวิชาออนไลน์และลิเนียร์ไอซีด้วยเทคโนโลยีสื่อสารทางไกลและเทคโนโลยีสารสนเทศที่พัฒนาขึ้น เสนอเป็นเอกสารเชิงแนวคิด เพื่อให้เกิดการวิพากษ์ในมุมมองจากผู้ทรงคุณวุฒิด้าน

เทคโนโลยีการศึกษา และสรุปเป็นรูปแบบการจัดการเรียนการสอน โดยการสัมมนากลุ่ม (Focus Group) และเพื่อตรวจสอบความตรงด้านเนื้อหาและความเหมาะสมของรูปแบบ แล้วนำมาแก้ไขปรับปรุงตามข้อเสนอแนะของผู้ทรงคุณวุฒิ

ประชากรศึกษา คือ ผู้ทรงคุณวุฒิด้านสื่อและเทคโนโลยีการศึกษา จำนวน 5 คน

เครื่องมือที่ใช้ คือ 1. แบบสังเกต แบบสัมภาษณ์แบบมีโครงสร้าง โดยผ่านการตรวจสอบเครื่องมือจากที่ปรึกษาโครงการวิจัย 2. รูปแบบการจัดการเรียนการสอนด้วยเทคโนโลยีสื่อสารทางไกลและเทคโนโลยีสารสนเทศ แบบประเมินรูปแบบการจัดการเรียนการสอนด้วยเทคโนโลยีสื่อสารทางไกลและเทคโนโลยีสารสนเทศ

สถิติที่ใช้ในการวิจัย คือ ค่าเฉลี่ย \bar{x} และ ค่าส่วนเบี่ยงเบนมาตรฐาน (S.D.)

4. ผลการดำเนินงาน

ผลการวิจัยแสดงได้ดังนี้ 4.1 รูปแบบการจัดการเรียนการสอนรายวิชาออนไลน์และลิเนียร์ไอซีด้วยเทคโนโลยีสื่อสารทางไกลและเทคโนโลยีสารสนเทศ HCL Model (Hyflex Constructionism Model) เป็นรูปแบบการจัดการเรียนการสอนแบบยืดหยุ่นตามสภาพแวดล้อมการเรียนรู้ ใช้กระบวนการจัดการเรียนการสอนที่ให้ผู้เรียนสร้างองค์ความรู้ด้วยตนเอง ซึ่งผู้เรียนสามารถเรียนได้ทั้งในห้องเรียน นอกห้องเรียน ทุกที่ ทุกเวลา ประกอบด้วยปัจจัยหลัก 4 ประการ คือ 1. ปัจจัยนำเข้า 2. ปัจจัยด้านกระบวนการ 3. ปัจจัยนำออก และ 4. ปัจจัยย้อนกลับ ประกอบด้วยขั้นตอนการจัดการเรียนการสอน 11 ขั้นตอนหลัก และ 23 ขั้นตอนย่อย คือ 1. ขั้นการวิเคราะห์ 1.1) วิเคราะห์หลักสูตร 1.2) วิเคราะห์เนื้อหารายวิชา 1.3) วิเคราะห์ผู้เรียน 1.4) วิเคราะห์ระบบพื้นฐาน 2. ขั้นการออกแบบและพัฒนาระบบ 2.1) กำหนดเป้าหมายการเรียนรู้ 2.2) พัฒนาวัตถุประสงค์การสอน 2.3) รวบรวมและออกแบบสร้างสรรค์เนื้อหา 2.4) เลือกกิจกรรมการสอนให้ผู้เรียนมีส่วนร่วมในการเรียนรู้ 2.5) สร้างสรรค์วิธีการสื่อสารเพื่อการสอน 2.6) เลือกวิธีการถ่ายทอดเนื้อหา 2.7) เตรียมแหล่งสนับสนุนการเรียนรู้และแหล่งข้อมูลอ้างอิงที่เกี่ยวข้อง 3. ขั้นการเตรียมการ 3.1) การปฐมนิเทศ 3.2) รับทราบเป้าหมายหรือวัตถุประสงค์ 3.3) ขอบข่าย 4. ขั้นแสวงหาความรู้ 5. ขั้นรวบรวมและจัดการกับข้อมูลที่เกี่ยวข้อง 6.

ขั้นวางแผนสร้างชิ้นงาน 7. ขั้นปฏิบัติการสร้างชิ้นงาน 7.1) การสืบค้น 7.2) การทดลองทำ 7.3) ผลจากการทดลอง 7.4) การตรวจสอบย้อนกลับดู 8. ขั้นการประเมินผล 8.1) ด้านความคืบหน้า 8.2) ด้านกระบวนการ 8.3) ด้านผลงาน 9) ขั้นนำเสนอผลงาน 10) ขั้นเผยแพร่ผลงาน 11) ขั้นการตรวจสอบ 11.1) ด้านความคืบหน้า และ 11.2) ด้านกระบวนการ

4.2 ประสิทธิภาพของรูปแบบที่พัฒนาขึ้น ผ่านการประเมินจากผู้ทรงคุณวุฒิด้านสื่อและเทคโนโลยีการศึกษาจำนวน 5 ท่าน มีความเห็นว่ารูปแบบการจัดการเรียนการสอนรายวิชา ออปแอมป์และลิเนียร์ไอซี ด้วยเทคโนโลยีสื่อสารทางไกลและเทคโนโลยีสารสนเทศมีคุณภาพและเหมาะสมของปัจจัยทั้ง 4 ด้านและตามลำดับขั้นตอนต่างๆ โดยรวมในระดับมากที่สุด (ค่าเฉลี่ย $\bar{x} = 4.75$, S.D. = 0.07) เมื่อแยกพิจารณาเป็นรายด้าน ผู้ทรงคุณวุฒิมีความเห็นว่าปัจจัยด้านกระบวนการ มีความเหมาะสมมากที่สุดที่ค่าเฉลี่ย $\bar{x} = 4.82$, S.D. = 0.030 รองลงมาคือปัจจัยนำออกค่าเฉลี่ย $\bar{x} = 4.80$, S.D. = 0.000 ปัจจัยย้อนกลับค่าเฉลี่ย $\bar{x} = 4.70$, S.D. = 0.140 และปัจจัยนำเข้า ค่าเฉลี่ย $\bar{x} = 4.67$, S.D. = 0.030

5. สรุป

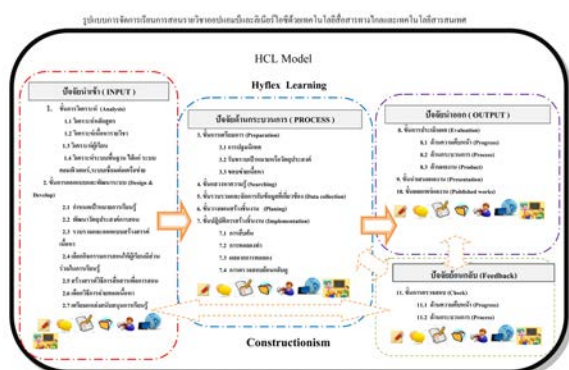
การวิจัยครั้งนี้ทำให้ได้รูปแบบการจัดการเรียนการสอนรายวิชาออปแอมป์และลิเนียร์ไอซีด้วยเทคโนโลยีสื่อสารทางไกลและเทคโนโลยีสารสนเทศ ที่เรียกว่า HCL Model (Hyflex Constructionism Learning) ที่สามารถนำไปใช้ในการเรียนการสอนในศตวรรษที่ 21 สำหรับผู้เรียนที่มีความสนใจการเรียนผ่านการใช้เทคโนโลยีซึ่งสามารถเรียนที่ไหนก็ได้ตามต้องการทั้งในห้องเรียน นอกห้องเรียน ทุกที่ ทุกเวลา นำไปสู่การเรียนรู้ตลอดชีวิต (Long Life) โดยใช้กระบวนการจัดการเรียนการสอนที่ให้ผู้เรียนสร้างองค์ความรู้ด้วยตนเอง ประกอบด้วย องค์ประกอบหลัก 4 ประการ คือ 1. ปัจจัยนำเข้า (INPUT) 2. ปัจจัยด้านกระบวนการ (PROCESS) 3. ปัจจัยนำออก (OUTPUT) และ 4. ปัจจัยย้อนกลับ (FEEDBACK) ประกอบด้วย ขั้นตอนการจัดการเรียนการสอน 11 ขั้นตอนหลักและ 22 ขั้นตอนย่อย คือ 1. ขั้นการวิเคราะห์ (Analysis) เป็นขั้นตอนที่ผู้สอนทำการวิเคราะห์ปัจจัยที่จำเป็นและเกี่ยวข้องในการจัดการเรียนรู้ที่ต้องใช้สื่อประสมทั้งในห้องเรียนและนอกห้องเรียน ประกอบด้วยขั้นตอนย่อยคือ 1.1) วิเคราะห์หลักสูตร 1.2)

วิเคราะห์เนื้อหารายวิชา 1.3) วิเคราะห์ผู้เรียน 1.4) วิเคราะห์ระบบพื้นฐาน 2. ขั้นการออกแบบและพัฒนาระบบ (Design & Development) เป็นขั้นตอนที่ผู้สอนทำการออกแบบการจัดการเรียนการสอนที่ใช้สื่อประสมที่มีความหลากหลายในการเรียน ทั้งการเรียนในห้องเรียนและการเรียนนอกห้องเรียน โดยการใช้เทคโนโลยีซึ่งประกอบด้วยขั้นตอนย่อยคือ 2.1) กำหนดเป้าหมายการเรียนรู้ 2.2) พัฒนาวัตถุประสงค์การสอน 2.3) รวบรวมและออกแบบสร้างสรรค์เนื้อหา 2.4) เลือกกิจกรรมการสอนให้ผู้เรียนมีส่วนร่วมในการเรียนรู้ 2.5) สร้างสรรค์วิธีการสื่อสารเพื่อการสอน 2.6) เลือกวิธีการถ่ายทอดเนื้อหา 2.7) เตรียมแหล่งสนับสนุนการเรียนรู้และแหล่งข้อมูลอ้างอิงที่เกี่ยวข้อง 3. ขั้นการเตรียมการ (Preparation) เป็นขั้นตอนก่อนเริ่มการเรียนการสอนประกอบด้วยขั้นตอนย่อยคือ 3.1) การปฐมนิเทศ 3.2) รับประทานอาหารหรือวัตถุประสงค์ 3.3) ขอบข่าย 4. ขั้นแสวงหาความรู้ (Searching) เป็นขั้นตอนขณะการจัดการเรียนการสอนโดยการค้นคว้าหาข้อมูลที่เกี่ยวข้องจากแหล่งเรียนรู้ต่างๆ 5. ขั้นรวบรวมและจัดการกับข้อมูลที่เกี่ยวข้อง (Data Collection) เป็นขั้นตอนที่ผู้เรียนรวบรวมและเลือกข้อมูลที่เกี่ยวข้องที่สามารถอ้างอิงได้ถูกต้อง 6. ขั้นวางแผนสร้างชิ้นงาน (Planning) เป็นขั้นตอนของการนำข้อมูลที่ได้นำมาวางแผนสร้างผลงาน ชิ้นงานว่ามีความเหมาะสมหรือไม่อย่างไร 7. ขั้นปฏิบัติการสร้างชิ้นงาน (Implementation) เป็นขั้นตอนการปฏิบัติการทดลองสร้างชิ้นงาน โดยมีขั้นตอนย่อยคือ 7.1) การสืบค้น เป็นการสืบค้นข้อมูลที่เกี่ยวข้องในสการปฏิบัติการทดลองเพิ่มเติม 7.2) การทดลองทำ เป็นการปฏิบัติการทดลอง 7.3) ผลจากการทดลอง เป็นผลการที่ได้จากการปฏิบัติการทดลอง 7.4) การตรวจสอบย้อนกลับดู เป็นขั้นตอนตรวจสอบความถูกต้องของการทดลอง ผลการทดลอง และแก้ไขผลงาน/ชิ้นงานให้มีความสมบูรณ์มากยิ่งขึ้น 8. ขั้นการประเมินผล (Evaluation) เป็นขั้นตอนของการประเมินผล ผลงาน/ชิ้นงาน ประกอบด้วยขั้นตอนย่อยคือ 8.1) ด้านความคืบหน้า (Progress) เป็นขั้นตอนการประเมินผลความคืบหน้าของการจัดทำผลงาน/ชิ้นงาน 8.2) ด้านกระบวนการ (Process) เป็นขั้นตอนการประเมินผลกระบวนการจัดทำผลงาน/ชิ้นงาน 8.3) ด้านผลงาน (Product) เป็นการประเมินผลของผลงาน/ชิ้นงาน 9) ขั้นนำเสนอผลงาน (Presentation) เป็นขั้นตอนของการนำเสนอ

ผลงาน/ชิ้นงาน10)ขั้นเผยแพร่ผลงาน (Published Works)เป็นขั้นตอนการเผยแพร่ผลงาน11)ขั้นการตรวจสอบ(Check)เป็นขั้นตอนการตรวจสอบของกระบวนการจัดการเรียนการสอนขององค์ประกอบหลัก 3 ประการคือปัจจัยนำเข้า ปัจจัยด้านกระบวนการ และปัจจัยนำออกประกอบด้วยขั้นตอนการตรวจสอบย่อยคือ111.)ด้านความคืบหน้า(Progress) เป็นการตรวจสอบด้านความคืบหน้าของผลงาน/ชิ้นงานและ112.)ด้านกระบวนการ (Process)เป็นขั้นตอนการตรวจสอบกระบวนการทำงานในการผลิตผลงาน/ชิ้นงาน

6. รูปภาพประกอบ

รูปแบบการจัดการเรียนการสอนรายวิชาออปแอมป์และลิเนียร์ไอซีด้วยเทคโนโลยีสื่อสารทางไกลและเทคโนโลยีสารสนเทศ HCL Model (Hyflex Constructionism Model)



ภาพที่ 1: แสดงรูปแบบการจัดการเรียนการสอนรายวิชาออปแอมป์และลิเนียร์ไอซีด้วยเทคโนโลยีสื่อสารทางไกลและเทคโนโลยีสารสนเทศ HCL Model (Hyflex Constructionism Model)

7. เอกสารอ้างอิง

[1] สำนักงานคณะกรรมการการอาชีวศึกษา . (2555). รายงานประจำปี 2555. กลุ่มประเมินผล สำนักติดตามและประเมินผลอาชีวศึกษา. กรุงเทพฯ : กระทรวงศึกษาธิการ.

[2] ควงนภามกรนุรักษ์. (2554). อนาคตภาพการอาชีวศึกษาไทยในทศวรรษหน้า(พ.ศ. 2554 - พ.ศ.2564).วารสารบริหารการศึกษา มหาวิทยาลัยขอนแก่น, ปีที่ 7,ฉบับที่ 2, หน้า 139-145.

[3] Barbara Seel และ Zita Glasgow (1998) อังโน ถักรพงศ์ ชูแสงนิล. 2555. การออกแบบระบบการสอน ADDIE MODEL. การออกแบบระบบการสอนรายวิชาการวิเคราะห์ทรัพยากรสารสนเทศ.

[ออนไลน์]. เข้าถึงได้จาก : <http://codexarticle.blogspot.com./2012/04/addie-model.html>. (วันที่ค้นข้อมูล : 10 เมษายน 2555).

[4] Brian J. Beatty, (2012),Choose HyFlex your alternative. InstructionalTechnologies San Francisco State University, (online), <http://itec.sfsu.edu/hyflex/hyflexhome.htm>2011

[5] ทิศนา แจมมณี. (2551). รูปแบบการเรียนการสอนทางเลือกที่หลากหลาย. พิมพ์ครั้งที่ 5 (ฉบับพิมพ์เพิ่มเติม) กรุงเทพฯ : สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย. มุลนิธิสื่อเพื่อการศึกษาของชุมชน, 2554, หน้า 147

[6] อุทิศ บำรุงชีพ. (2551). รูปแบบการเรียนการสอนผ่านเว็บตามแนวคอนสตรัคชันนิซึมเพื่อพัฒนาความคิดสร้างสรรค์. ปรัญญาศึกษาศาสตร์ดุสิตบัณฑิต (เทคโนโลยีการศึกษา) สาขาวิชาเทคโนโลยีการศึกษา ภาควิชาเทคโนโลยีการศึกษา มหาวิทยาลัยเกษตรศาสตร์.

ระบบสารสนเทศการจัดการการเรียนการสอน Learning Management Information System

ปานวิทย์ ฐะนุติ (Panwit Tuwanut)¹, ศักดิ์สิทธิ์ สว่าง (Saksit Sawang)² และ ณิชกมล ศรีภูมิ (Nutmamol Sripum)³

คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

Emails: ¹panwit@it.kmitl.ac.th, ²hanamiza555@gmail.com, ³farina_liszt@hotmail.com

บทคัดย่อ

เนื่องจากในปัจจุบัน การเรียนการสอนในระดับต่างๆ ได้มีการนำเทคโนโลยีสารสนเทศต่างๆ มาประยุกต์ใช้ในการเรียนการสอน เป็นจำนวนมาก แต่ระบบที่พบในปัจจุบันยังขาดการเชื่อมโยงระหว่างสถานศึกษา กับ ผู้ปกครอง เข้าด้วยกัน ดังนั้นในงานการพัฒนากระบวนนี้จึงเป็นการนำเสนอ การพัฒนาระบบสารสนเทศเพื่อจัดการ การเรียนการสอน โดยมีวัตถุประสงค์เพื่อเชื่อมโยง อาจารย์ นักศึกษา และ ผู้ปกครอง เข้าด้วยกัน โดยพัฒนาระบบเป็น 2 ส่วน คือ เว็บแอปพลิเคชัน พัฒนาด้วยภาษา Python และ แอปพลิเคชันบนไอแพด พัฒนาด้วยภาษา Objective-C และ ใช้ระบบจัดการฐานข้อมูล MySQL

คำสำคัญ— การเรียนการสอน การศึกษา iOS

Abstract

Nowadays of teaching in classroom. Have adopted various technology to applied in classroom. But in the current system still lacked the connection between academy and parent together.

In this system development, we developed a system that will manage classroom teaching. That will connecting teacher, student and parent together. The system has 2 parts. First, a web application and it's developing with Python language. Second, a mobile application on iPad and it's developing in Objective-C language along with MySQL database.

Keywords: Learning Management, Education, iOS

1. บทนำ

เนื่องจากปัจจุบันในการเตรียมการสอนแต่ละครั้ง อาจารย์จำเป็นต้องเตรียมเอกสารต่างๆ เป็นจำนวนมาก อาทิเช่น เอกสารประกอบการสอน เอกสารในการเช็คชื่อ รวมถึงแบบทดสอบต่างๆ โดยที่ผ่านมาได้มีการพัฒนาระบบสารสนเทศต่างๆ ที่อำนวยความสะดวกในการเรียนการสอน อาทิเช่น ระบบห้องเรียนเสมือนจริง ของ สุรวุฒิ อนุรักษย์, สุวิทย์ สมส่งกุล [1] และ การพัฒนาระบบการสอนแบบห้องเรียนเสมือนบนเครือข่ายคอมพิวเตอร์ ของ สุวิทย์ อิบมันตะศิริ [2] ซึ่งเน้นไปที่การเรียนการสอนเสมือนได้เรียนรู้อยู่ในห้องเรียน นอกจากนี้ยังมีระบบ E-learning ที่พัฒนาโดยการใช้โปรแกรม Moodle ซึ่งเป็น ระบบจัดการเรียนการสอนในระบบออนไลน์ สำหรับอาจารย์ ใช้เพื่อเตรียมแหล่งข้อมูล กิจกรรม และเผยแพร่แบบออนไลน์ผ่านอินเทอร์เน็ต หรืออินทราเน็ต [3] ซึ่งระบบทั้งหมดดังกล่าวมาข้างต้นมีข้อจำกัดคือ จะมุ่งเน้นไปที่อาจารย์ และนักศึกษาเท่านั้น ดังนั้นในงานพัฒนากระบวนนี้จึงได้พัฒนาระบบสารสนเทศการจัดการการเรียนการสอน เพื่อมุ่งเน้นไปที่ การเชื่อมโยงระหว่าง สถาบันการศึกษา และ ผู้ปกครองเข้าด้วยกัน เพื่อให้ผู้ปกครองช่วยสอดส่องพฤติกรรมของนักศึกษา และ ระบบนี้ช่วยทำให้ผู้ปกครองสามารถรับรู้พฤติกรรมการเรียนต่างๆของนักศึกษาได้ และ อำนวยความสะดวกในด้านการเรียนการสอนระหว่างอาจารย์และนักศึกษา โดยพัฒนาระบบเป็นเว็บแอปพลิเคชันและแอปพลิเคชันบนไอแพด เพื่อให้ผู้ใช้งานสามารถใช้งานได้สะดวกยิ่งขึ้น

โดยเนื้อหาของบทความส่วนที่ 2 ได้กล่าวถึงทฤษฎีและหลักการที่เกี่ยวข้องในการพัฒนาระบบ ส่วนที่ 3 กล่าวถึงการวิเคราะห์และออกแบบระบบ ส่วนที่ 4 กล่าวถึงระบบต้นแบบ และส่วนที่ 5 เป็นสรุปผล

2. ทฤษฎีและหลักการที่เกี่ยวข้อง

2.1 ภาษา Python

ภาษา Python เป็นภาษาที่มีการประมวลผลแบบอินเทอร์พรีเตอร์ (Interpreted programming language) [4] คือจะประมวลผลและปฏิบัติตามคำสั่งไปที่ละบรรทัด ผู้คิดค้นคือ Guido Van Rossum ในปี ค.ศ. 1990 ซึ่งภาษา Python เป็นการจัดการชนิดของตัวแปรแบบแปรผันตามข้อมูลที่บรรจุอยู่ (Fully dynamically typed) และใช้การจัดการหน่วยความจำเป็นอัตโนมัติ (Automatic memory management) โดยเป็นการพัฒนาและผสมผสานของภาษาอื่น ๆ ได้แก่ ABC, Modula-3, Icon, ANSI C, Perl, Lisp, Smalltalk และ Tcl นอกจากนี้ ภาษา Python ยังเป็นแนวคิดที่ทำให้เกิดภาษาใหม่ ๆ ซึ่งได้แก่ Ruby และ Boo เป็นต้น

2.2 ภาษา Objective-C

ภาษา Objective-C เป็นภาษาที่พัฒนามาตั้งแต่ปี ค.ศ. 1986 โดย Brad Cox และ Tom Love โดยเป็นการพัฒนาต่อออกมาจากภาษา C และนำแนวคิดเชิงวัตถุ (Object-Oriented Concept) มาใช้และปรับปรุงให้ดีขึ้น โดยภาษา Objective-C เป็นภาษาที่สามารถ พัฒนาแอปพลิเคชันที่ใช้งานบนระบบปฏิบัติการ OSX และ iOS ได้ โดยมีโปรแกรมเอ็กซ์โค้ด (Xcode) ซึ่งเป็นโปรแกรมที่สามารถใช้งานได้บนระบบปฏิบัติการ OSX เท่านั้น เป็นเครื่องมือที่ใช้ในการพัฒนา

2.3 สถาปัตยกรรม MVC

MVC (Model View Controller) เป็นสถาปัตยกรรมซอฟต์แวร์ (Software Architecture) ที่มีการแบ่งแยกระบบออกเป็น 3 ส่วนหลักๆตามบทบาทหน้าที่ ได้แก่ Model, View, Controller [5]

2.3.1 ส่วนประกอบของ MVC

1. Model คือ ส่วนที่ตัดสินใจในการเข้าถึงและใช้งานข้อมูล โดยจะเป็นส่วนที่ทำงานติดต่อกับฐานข้อมูล (Database) จัดการข้อมูลเข้า-ออก เพื่อนำไปประมวลผล

2. View คือ ส่วนที่ทำหน้าที่ในการแสดงผล โดยจะนำข้อมูลที่ได้จาก Model มาแสดงผลผ่านทางเว็บเบราว์เซอร์ (Web Browser) ในรูปแบบของ HTML ให้ผู้ใช้ได้รับทราบข้อมูลต่างๆ

3. Controller คือ ส่วนของการประมวลผลหลักของระบบ จะคอยตอบรับและโต้ตอบการทำงานของผู้ใช้ (User Action) จากไคลเอ็นท์ (Client) คอยรับคำสั่งมาประมวลผล และส่งต่อไปให้ Model หรือ View ทำงานต่อไป

2.3.2 หลักการทำงานของ MVC

1. เมื่อไคลเอ็นท์ (Client) มีการร้องขอผ่าน View
2. ข้อมูลจะถูกส่งผ่านไปหา Controller โดยส่วน Controller จะเป็นตัวดำเนินการทำงาน (Action) ต่างๆ หรือเป็นส่วนของการประมวลผลของระบบ
3. จากนั้น Controller ก็จะร้องขอไปยัง Model จากการทำงาน (Action) นั้นๆ เช่น ร้องขอการเลือกข้อมูลในตารางทั้งหมด
4. เมื่อ Model รับการร้องขอก็จะทำการ Query ข้อมูลตามที่ส่วน Controller ส่งมา
5. Model จะส่งข้อมูลที่ได้กลับมาจากส่วน Controller
6. Controller ก็จะทำการกำหนดค่าลงในตัวแปรเพื่อส่งให้ส่วนของ View ต่อไป
7. ส่วนของ View ก็จะนำตัวแปรเหล่านั้นไปทำการแสดงผลตามต้องการได้ [6]

2.4 SQL

SQL (Structured Query Language) คือภาษาที่ใช้ในการเขียนโปรแกรม เพื่อจัดการกับฐานข้อมูล เป็นภาษามาตรฐานบนระบบฐานข้อมูลเชิงสัมพันธ์และเป็นระบบเปิด (open system) หมายถึงเราสามารถใส่คำสั่ง SQL กับฐานข้อมูลชนิดใดก็ได้ และ คำสั่งงานเดียวกันเมื่อสั่งงานผ่านระบบฐานข้อมูลที่แตกต่างกันจะได้ผลลัพธ์เหมือนกัน ทำให้เราสามารถเลือกใช้งานฐานข้อมูล ชนิดใดก็ได้โดยไม่ติดขัดกับฐานข้อมูลใดฐานข้อมูลหนึ่ง [7]

2.5 Google Drive

เป็นการเก็บข้อมูลในรูปแบบของ Cloud Storage หรือบริการจัดเก็บไฟล์ข้อมูลไว้บนเซิร์ฟเวอร์ ซึ่งทำให้ผู้ใช้งานสามารถเปิดใช้ไฟล์ข้อมูลต่างๆเหล่านั้นได้ทุกที่ทุกเวลา[8] และไม่ว่าจะใช้คอมพิวเตอร์ Notebook, PC มือถือ หรือ แท็บเล็ต ก็สามารถเข้าถึงไฟล์ข้อมูลต่างๆได้อย่างง่ายดายนอกจากนั้น Google Drive ยังเชื่อมโดยตรงกับบริการ Google Docs ซึ่ง

เป็นบริการจัดการเอกสารแบบออนไลน์ ทำให้สามารถแก้ไขไฟล์งาน เช่น ไฟล์เอกสาร หรือ ฟรีเซนเทชัน ร่วมกับคนอื่นๆ ได้แบบ real-time

3. การวิเคราะห์และออกแบบระบบ

3.1. การเก็บรวบรวมความต้องการจากผู้ใช้งาน

จากการสอบถามผู้ปกครองจำนวน 30 คน พบว่าผู้ปกครองมีความต้องการทราบข้อมูลต่างๆดังนี้

- | | |
|---------------------------|------|
| 1. คะแนนสอบ | 18 % |
| 2. ผลการเรียน | 17 % |
| 3. ข้อมูลการเรียน | 16 % |
| 4. ตารางเรียน | 15 % |
| 5. ความประพฤติในห้องเรียน | 12 % |
| 6. ข้อมูลรายวิชา | 12 % |
| 7. ตารางสอบ | 10 % |

3.2 ขอบเขตของระบบสารสนเทศ

3.2.1 ระบบทำงานร่วมกับผู้ใช้งาน 3 กลุ่ม ได้แก่

1. อาจารย์ สามารถอัปโหลดเอกสารประกอบการสอน เพื่อให้ให้นักศึกษาใช้ประกอบการเรียน เช็คชื่อ ให้คะแนนความประพฤติ ประกาศคะแนน กำหนดสัดส่วนคะแนน ตัดเกรด ประกาศข่าวต่างๆ ถาม-ตอบคำถาม และสร้างแบบทดสอบ

2. นักศึกษา สามารถเข้าไปดาวน์โหลดเอกสารประกอบการเรียน ดูประกาศข่าวต่างๆ ถาม-ตอบคำถาม ทำแบบทดสอบและดูคะแนนสอบของตนเองได้

3. ผู้ปกครอง สามารถตรวจสอบการเข้าเรียนของนักศึกษา สามารถตรวจสอบคะแนนต่างๆได้ และมีระบบแจ้งเตือนในกรณีต่างๆ เช่น นักศึกษาขาดเรียน หรือ คะแนนสอบได้น้อย เป็นต้น

3.2.2 แอปพลิเคชัน แบ่งเป็น 2 ส่วนได้แก่

1. แอปพลิเคชันบนไอแพดพัฒนาด้วยภาษา Objective-C สำหรับอาจารย์ผู้สอน และนักศึกษา
2. เว็บแอปพลิเคชันพัฒนาด้วยภาษา Python (web.py) สำหรับอาจารย์ นักศึกษา และผู้ปกครอง

3.2.3 ระบบฐานข้อมูล

1. ออกแบบฐานข้อมูลโดยใช้ SQL Object (MySQL)

3.3 การวิเคราะห์ความต้องการของระบบ

ความต้องการที่เป็นหน้าที่หลักของระบบ (Functional Requirement)

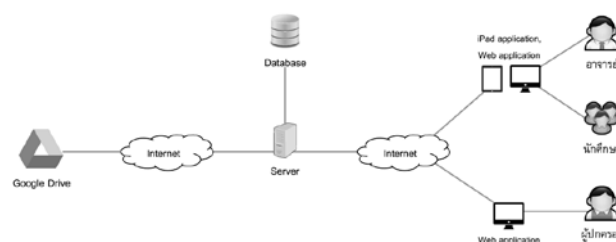
1. ระบบจะเก็บข้อมูลของผู้ปกครองในครั้งแรกที่ลงทะเบียน
2. อาจารย์สามารถเช็คชื่อของนักศึกษาได้
3. อาจารย์สามารถจัดการคะแนนของนักศึกษาได้
4. อาจารย์สามารถอัปโหลดเอกสารประกอบการเรียนได้
5. อาจารย์สามารถจัดการแบบทดสอบต่างๆได้
6. อาจารย์สามารถประกาศข่าวสารและถาม-ตอบคำถามต่างๆได้
7. นักศึกษาสามารถดาวน์โหลดเอกสารประกอบการเรียนได้
8. นักศึกษาสามารถดูผลคะแนนต่างๆได้
9. นักศึกษาสามารถทำแบบทดสอบต่างๆได้
10. นักศึกษาสามารถถาม-ตอบคำถามต่างๆได้
11. ผู้ปกครองสามารถดูพฤติกรรมการเรียนของนักศึกษาได้

ความต้องการที่ไม่ใช่หน้าที่หลักของระบบ (Non-Functional Requirement)

1. ระบบสามารถประมวลผลได้รวดเร็ว
2. ระบบมีส่วนติดต่อผู้ใช้งานที่สวยงาม
3. ระบบมีความปลอดภัยในการใช้งาน

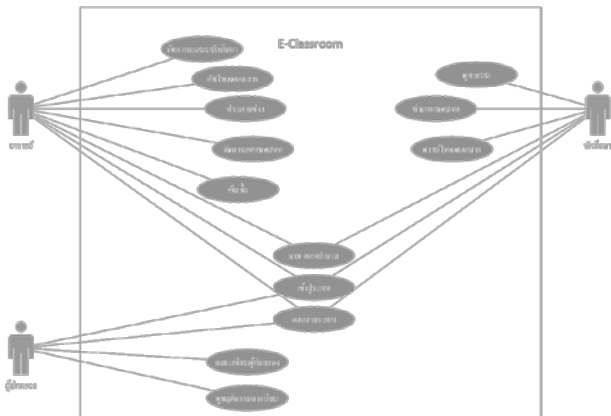
3.4 บล็อกไดอะแกรมแสดงภาพรวมของระบบ

แสดงภาพรวมของระบบสารสนเทศการจัดการการเรียนการสอน โดยระบบจะแบ่งเป็น 2 ส่วน คือ แอปพลิเคชันบนไอแพดและเว็บแอปพลิเคชัน มีการเก็บข้อมูลลงในฐานข้อมูลประเภท SQL และสำหรับเอกสารประกอบการสอนของอาจารย์จะเก็บไว้ในโฟลเดอร์ใน Google Drive และระบบจะแชร์โฟลเดอร์เหล่านั้นไปยังนักศึกษาที่ลงทะเบียนวิชานั้นๆ โดยอัตโนมัติ



ภาพที่ 1: แผนภาพบล็อกไดอะแกรม (Block Diagram)

3.5 แผนภาพยูสเคส (Use Case Diagram)



ภาพที่ 2: แผนภาพยูสเคส (Use Case Diagram)

เป็นแผนภาพที่แสดงความสัมพันธ์ระหว่าง Use Case และ Actors ที่มีการทำงานที่เชื่อมโยงกันโดยมีรายละเอียดดังนี้

3.5.1 ผู้เกี่ยวข้องกับระบบ (Actor)

ประกอบด้วย นักศึกษา (Student) อาจารย์ (Teacher) และ ผู้ปกครอง (Parent)

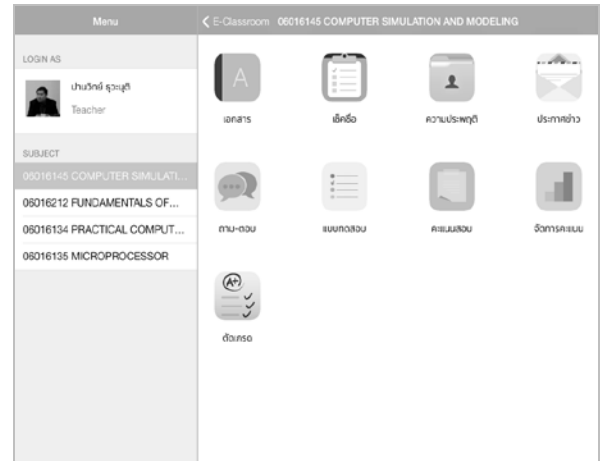
3.5.2 องค์ประกอบของแผนภาพยูสเคส

ผู้ใช้งานระบบสารสนเทศการจัดการการเรียนการสอน ประกอบด้วย อาจารย์ (Teacher) นักศึกษา (Student) และ ผู้ปกครอง (Parent)

1. เข้าสู่ระบบ
2. เช็คชื่อ
3. จัดการแบบทดสอบ
4. ประกาศข่าว
5. ถาม-ตอบคำถาม
6. จัดการคะแนนนักศึกษา
7. อัปโหลดเอกสาร
8. ดาวน์โหลดเอกสาร
9. ทำแบบทดสอบ
10. ดูคะแนน
11. ลงทะเบียนผู้ปกครอง
12. ดูพฤติกรรมกรรมการเรียน
13. ออกจากระบบ

4. ระบบต้นแบบ

4.1 ส่วนของอาจารย์

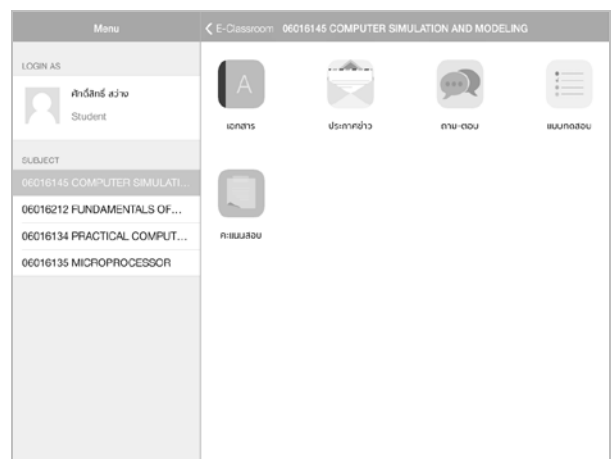


ภาพที่ 3: หน้าหลักของรายวิชาส่วนของอาจารย์

ประกอบด้วยเมนูต่างๆ โดยแบ่งเป็น 9 เมนูคือ

1. เอกสาร – สำหรับเปิดดูเอกสารที่ใช้ประกอบการสอน
2. เช็คชื่อ – สำหรับเช็คชื่อนักศึกษาในแต่ละวิชา
3. ความประพฤติ – สำหรับให้คะแนนความประพฤติในห้องของนักศึกษา
4. ประกาศข่าว – สำหรับประกาศข่าวต่างๆให้นักศึกษารับรู้
5. ถาม-ตอบ – สำหรับถามหรือตอบคำถามกับนักศึกษา
6. แบบทดสอบ – สำหรับสร้างแบบทดสอบต่างๆ
7. คะแนนสอบ – สำหรับประกาศคะแนนสอบต่างๆ
8. จัดการคะแนน – สำหรับจัดการสัดส่วนคะแนนในแต่ละวิชา
9. ตัดเกรด – สำหรับตัดเกรดของนักศึกษา

4.2 ส่วนของนักศึกษา



ภาพที่ 4: หน้าหลักของรายวิชาส่วนของนักศึกษา

ประกอบด้วยเมนูต่างๆ โดยแบ่งเป็น 5 เมนูคือ

- เอกสาร ใช้สำหรับเปิดดูเอกสารประกอบการเรียน
- ประกาศข่าว ใช้สำหรับดูข่าวต่างๆ
- ถาม-ตอบ ใช้สำหรับถามหรือตอบคำถามกับอาจารย์
- แบบทดสอบ ใช้สำหรับทำแบบทดสอบต่างๆ
- คะแนนสอบ ใช้สำหรับดูคะแนนสอบต่างๆ

4.3 ส่วนของผู้ปกครอง



ภาพที่ 5: หน้าแสดงข้อมูลการเรียน

ประกอบด้วย 4 เมนูคือ

1. ข้อมูลรายวิชา – แสดงข้อมูลของรายวิชานั้นๆ เช่นเวลาเรียน ชื่ออาจารย์ผู้สอน
2. ความประพฤติ – แสดงความประพฤติในห้องเรียนของนักศึกษา
3. คะแนน – แสดงคะแนนสอบต่างๆของนักศึกษา
4. Conversation – สำหรับติดต่อกับอาจารย์ในแต่ละรายวิชา

5. สรุปผล

โครงการนี้เป็นการศึกษาออกแบบและพัฒนาระบบที่สามารถเชื่อมโยงผู้ปกครอง อาจารย์และนักศึกษาเข้าด้วยกัน และช่วยอำนวยความสะดวกในด้านการเรียนการสอนระหว่างอาจารย์และนักศึกษา โดยพัฒนาระบบเป็น 2 ส่วน คือ เว็บไซต์แอปพลิเคชันพัฒนาโดยใช้ภาษา Python และ แอปพลิเคชันบนไอแพด พัฒนาโดยใช้ภาษา Objective-C

ระบบทำงานร่วมกับผู้ใช้งาน 3 กลุ่มได้แก่ อาจารย์ นักศึกษา ผู้ปกครอง โดยอาจารย์จะสามารถอัปโหลดเอกสารประกอบการสอน เช็กรายชื่อ จัดการคะแนนต่าง ประกาศข่าวต่างๆ ถาม-ตอบคำถาม และสร้างแบบทดสอบได้ นักศึกษาสามารถเข้าไปดูแนวโน้มผลเอกสารประกอบการเรียน ดูประกาศข่าวต่างๆ ถาม-ตอบคำถาม ทำแบบทดสอบและดูคะแนนสอบของ

ตนเองได้ ผู้ปกครองสามารถตรวจสอบการเข้าเรียนของนักศึกษาและสามารถตรวจสอบคะแนนต่างๆได้ โดยข้อมูลจะจัดเก็บในรูปแบบของฐานข้อมูลแบบ SQL

เอกสารอ้างอิง

- [1] สุวิทย์ สมส่งกุล และ สุรวุฒิ อนุรักษ์. 2547 “ระบบห้องเรียนเสมือนจริง” ปริญญาโท. คณะวิศวกรรมศาสตร์ สาขาวิชาวิศวกรรมคอมพิวเตอร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
- [2] สุวิทย์ ยิบมันตะศิริ. “การพัฒนาระบบการสอนแบบห้องเรียนเสมือนบนเครือข่ายคอมพิวเตอร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.” วิทยานิพนธ์ครุศาสตรบัณฑิต สาขาวิชาวิศวกรรมไฟฟ้าสื่อสาร บัณฑิตวิทยาลัย,สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง. 2546
- [3] “Moodle คืออะไร.” [Online] Availablefrom : [http:// dek-ac.com/d-sign2u/cms-id7.html](http://dek-ac.com/d-sign2u/cms-id7.html). 2013
- [4] “Python.” [Online] Availablefrom : <http://accdekdeepython.blogspot.com/2011/02/python-python-python-open-source-python.html>. 2013
- [5] “MVC คืออะไร.” [Online] Availablefrom : <http://dotnetoverthailand.blogspot.com/2013/07/v-behaviorurldefaultvml.html>. 2013
- [6] “หลักการทํางานของ MVC.” [Online] Availablefrom : <https://code.google.com/p/mollow/wiki/MVC>. 2013
- [7] “SQL.” [Online] Availablefrom : <http://www.softwaresiam.com/index.php/sql/11-sql>. 2013
- [8] Peerapat Chuejeen, “5 ข้อดีของ Google Drive ที่คุณอาจยังไม่รู้.” [Online] Availablefrom : <http://www.techshabu.com/2013/10/16/5-benefit-of-google-drive/>. 2014

การพัฒนาการจัดการเรียนการสอนแบบออนไลน์สำหรับอาจารย์และนักศึกษา

มหาวิทยาลัยราชภัฏนครราชสีมา

The Development of the LMS Online for Teacher and Student in Nakhon Ratchasima Rajabhat University

สรายุทธ เนียนกระโทก (Sarayut Neankratok)¹ และ อุษานาฏ เอื้ออภิสิทธิ์วงศ์ (Usanad Ua-apisitwong)²

¹โปรแกรมวิชาวิทยาการสารสนเทศ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครราชสีมา

²โปรแกรมวิชาคอมพิวเตอร์ศึกษา คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครราชสีมา

¹yutsak_2005@hotmail.com, ²usanad@hotmail.com

บทคัดย่อ

การเรียนการสอนระดับมหาวิทยาลัยมีรูปแบบที่แตกต่างกันไปในแต่ละสถาบันการศึกษา คณะหรือรายวิชา เช่น การจัดรูปแบบการเรียนการสอน การจัดเก็บคะแนน การส่งงาน เป็นต้น ซึ่งมหาวิทยาลัยราชภัฏนครราชสีมาเป็นสถาบันอุดมศึกษาเพื่อท้องถิ่น จึงมีนักศึกษาเข้าเรียนในแต่ละปีเป็นจำนวนมากและมีความหลากหลาย การจัดการเรื่องของการให้คะแนนและการเรียนภายในห้องจึงไม่เพียงพอหรือครอบคลุมความเข้าใจของนักศึกษาได้ทั้งหมด วัตถุประสงค์ของการวิจัยคือการศึกษาและพัฒนาระบบการจัดการเรียนการสอนแบบออนไลน์ที่จะช่วยสนับสนุนการทำงานและการเรียนของครูและนักเรียน โดยข้อดีของระบบดังกล่าวคือใช้งานได้ง่ายและครอบคลุมความต้องการของผู้ใช้มากที่สุด วิธีการประเมินผลใช้การสัมภาษณ์ สอบถามและวิเคราะห์ระบบเดิม ผลการทดลองระดับความพอใจของอาจารย์อยู่ในระดับดี คิดเป็นค่าเฉลี่ย 4.01 และความพอใจของนักศึกษายู่ในระดับดี คิดเป็นค่าเฉลี่ย 4.34

คำสำคัญ: ระบบการเรียนการสอน บทเรียนออนไลน์ ความพึงพอใจ

Abstract

Teaching of an university has a different from each institution factory or subject such as an instruction detail Scoring or Assignment. So the Nakhon Ratchasima Rajabhat University is the local institution, then in each year the admission is a lot. The class management is insufficient and not comfortable for collecting the student score. The objective of this research is to study and

develop the web-based learning management system for supporting the working of a teacher or student. The advantage of this system is easy to use and covers the requirement of most user. The interview and asking are used to evaluate this system. The experimental results, the satisfaction level of the teacher is good (an average is 4.01) and the satisfaction level of the student is good (an average is 4.34).

Keywords: LMS, e-Learning, Satisfaction

1. บทนำ

ปัจจุบันรูปแบบการเรียนการสอนของสถาบันอุดมศึกษาในประเทศไทยมีการนำเอาบทเรียนออนไลน์ (E-learning) และระบบการจัดการเรียนการสอน (Learning Management System : LMS) มาใช้เป็นจำนวนมาก เพื่อช่วยในการเรียนการสอนให้มีประสิทธิภาพเพิ่มมากขึ้น เนื่องจากการจัดการเรียนการสอนผ่านเครือข่ายจะช่วยให้ผู้เรียนเกิดกระบวนการคิดวิเคราะห์ การศึกษาค้นคว้าจนเกิดเป็นองค์ความรู้ใหม่ๆ ได้เอง [1] และยังมีคามยืดหยุ่นในการจัดการเรียนการสอนในด้านต่างๆ เช่น เวลา คุณลักษณะผู้เรียนหรือบริบททางการเรียนรู้ เป็นต้น [2]

ระบบการจัดการเรียนการสอนถูกพัฒนาขึ้นเป็นจำนวนมาก โดยระบบจะมีความสามารถในการทำงานต่างกันและจากรูปแบบการเรียนการสอนในแต่ละสถาบันมีความแตกต่างทั้งเรื่องของวิธีการสอน ความสามารถในการเรียนรู้ ทรัพยากรและระบบการทำงานของมหาวิทยาลัย ทำให้ประสิทธิภาพในการใช้งานบทเรียนออนไลน์หรือระบบการจัดการเรียนการสอนที่มีอยู่เดิมลดลง เพราะไม่สอดคล้องกับความต้องการของผู้ใช้

นอกจากนี้ยังมีปัญหาในการใช้งานระบบการจัดการเรียนการสอน เช่น การยอมรับระบบหรือค่าใช้จ่าย เป็นต้น [2]

จากปัญหาด้านต่างๆ ผู้วิจัยได้พัฒนาระบบการจัดการเรียนการสอนแบบออนไลน์สำหรับอาจารย์และนักศึกษามหาวิทยาลัยราชภัฏนครราชสีมาเพื่อช่วยในการแก้ปัญหาการจัดการเรียนการสอนให้ครอบคลุมความสามารถในการเรียนรู้ การให้คะแนนและการรับส่งการบ้านของอาจารย์หรือนักศึกษาได้อย่างเต็มที่ ทำให้ประหยัดทรัพยากรและเวลาในการจัดเก็บเอกสารการสอน คะแนนเก็บ และการส่งงานของนักศึกษา นอกจากนี้รูปแบบของหน้าเว็บไม่ใช้ทรัพยากรของเครื่องคอมพิวเตอร์มากจึงช่วยในการประหยัดงบประมาณของการดูแลและการเข้าถึงระบบ จึงเป็นส่วนสำคัญในการช่วยเหลือให้การใช้งานระบบการเรียนการสอนมีประสิทธิภาพมากยิ่งขึ้น

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 บทเรียนออนไลน์

บทเรียนออนไลน์ หมายถึง การใช้เทคโนโลยีเพื่อสร้างหรือนำเสนอสื่อการเรียนการสอนในรูปแบบต่างๆ ผ่านเทคโนโลยีและการสื่อสาร (Information Communication Technology : ICT) [3] เช่น การเรียนการสอนผ่านอินเทอร์เน็ต หรือ ซีดีรอม เป็นต้น ตัวอย่างของสื่อการเรียนการสอนที่นิยมใช้ในปัจจุบันมีหลายแบบ เช่น การเรียนบนเว็บไซต์ (Web-based Instruction : WBI) คอมพิวเตอร์ช่วยสอน (Computer Assisted Instruction : CAI) การเรียนออนไลน์ (On-line Learning) หรือการเรียนทางไกลผ่านดาวเทียม เป็นต้น [4] องค์ประกอบของบทเรียนออนไลน์ประกอบไปด้วย 4 ส่วน คือ ระบบจัดการการศึกษา การสื่อสารระหว่างบุคคล ส่วนของเนื้อหาและการวัดผลการเรียน โดยประโยชน์ของบทเรียนออนไลน์มีอยู่หลายอย่างทั้งความยืดหยุ่นในการปรับเปลี่ยนเนื้อหา ความสะดวกในการเรียน เข้าถึงและ ปรับปรุงเนื้อหาใหม่ได้ง่าย ทำให้ประหยัดเวลาและค่าใช้จ่ายได้มาก [5]

2.2 การจัดการเรียนการสอน

การจัดการเรียนการสอน หมายถึง โปรแกรมการบริหารจัดการระบบการจัดการเรียนการสอนแบบบทเรียนออนไลน์เป็นระบบบริหารจัดการที่มีรูปแบบการทำงานที่มีความยืดหยุ่น และสามารถรองรับการจัดการเรียนการสอนได้หลายรูปแบบ

และมีประสิทธิภาพหรือสัมฤทธิ์ผลตามเป้าหมายที่กำหนดไว้ ซึ่งมุ่งเน้นที่การจัดการเกี่ยวกับนักศึกษาเป็นหลัก รวมถึงกิจกรรมของนักศึกษา การติดตามความก้าวหน้าและประเมินความสามารถของนักศึกษา [6]

2.3 งานวิจัยที่เกี่ยวข้อง

ฉมน จีรังสุวรรณ และ ธนยศ สิริโชค [7] ได้สร้างรูปแบบการเรียนการสอนแบบร่วมมือด้วยระบบการจัดการกิจกรรมการเรียนรู้ออกแบบเพื่อแก้ข้อบกพร่องกระบวนการเรียนการสอนให้มีความหลากหลายและยืดหยุ่นกับรูปแบบการเรียนในสภาพแวดล้อมปัจจุบัน ขั้นตอนวิธีการสร้างรูปแบบการเรียนการสอนแบบร่วมมือด้วยกระบวนการจัดการกิจกรรมการเรียนรู้ออกแบบมี 5 ขั้นตอน คือ

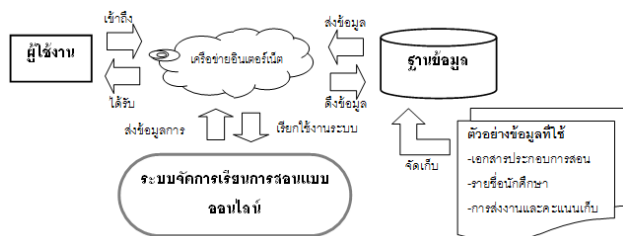
1. การศึกษาเอกสาร แนวความคิด ทฤษฎีและงานวิจัยที่เกี่ยวข้อง เพื่อทำการวิเคราะห์และสังเคราะห์ให้ได้ขั้นตอนการจัดรูปแบบที่สอดคล้องกับการเรียนการสอน
2. ร่างรูปแบบการเรียนการสอนแบบร่วมมือด้วยกระบวนการจัดการกิจกรรมการเรียนรู้ออกแบบได้จากข้อ 1.
3. สร้างระบบตามร่างรูปแบบและทำการตรวจสอบปรับปรุง
4. จัดส่งร่างรูปแบบการเรียนการสอนแบบร่วมมือด้วยกระบวนการจัดการกิจกรรมการเรียนรู้ออกแบบให้ผู้ช่วยชาญตรวจสอบความสอดคล้องขององค์ประกอบด้านปัจจัยนำเข้า กระบวนการควบคุมการผลิต และข้อมูลกลับ
5. ปรับปรุงรูปแบบการเรียนการสอนแบบร่วมมือด้วยกระบวนการจัดการกิจกรรมการเรียนรู้ออกแบบที่ผู้เชี่ยวชาญแนะนำก่อนนำไปใช้

พิจิต ตรีวิทย์รัตน์ และ วรณา ตรีวิทย์รัตน์ [8] ได้สร้างนวัตกรรมจัดการเรียนการสอนระบบอีเลิร์นนิ่งแบบจุดเดียวเบ็ดเสร็จ เพื่อลดปัญหาจากการพึ่งพาของผู้สอน ผู้ออกแบบสื่อการสอนและผู้บริหาร ซึ่งปัญหาที่เกิดจากระบบอีเลิร์นนิ่งในอดีต คือ ผู้สอนไม่เข้าใจในระบบการทำงานของอีเลิร์นนิ่ง ขาดความรู้และทักษะด้านเทคโนโลยี ขาดแคลนอุปกรณ์ทางคอมพิวเตอร์หรือผู้สอนมีภาระงานมาก เป็นต้น และปัญหาเรื่องของการจัดการเรียนการสอนหรือการไม่สนับสนุนงบประมาณ ขาดที่ปรึกษา นักศึกษาไม่สนใจศึกษาเพิ่มเติม เป็นต้น ซึ่งเป็นองค์ประกอบของความล้มเหลวของระบบอีเลิร์นนิ่ง ดังนั้น

จากจากปัญหาต่างๆ งานวิจัยนี้ทำการเพิ่มศักยภาพของผู้สอน การติดต่อประสานงาน การตรวจสอบระบบเครือข่าย การออกแบบสื่อการสอน การออกแบบบททดสอบ การสร้างสื่อการเรียน การจัดเตรียมผู้ช่วย เพื่อลดปัญหาการพึ่งพาบุคลากรด้านต่างๆ และเพื่อให้ระบบอิเล็กทรอนิกส์เป็นประโยชน์ เป็นทางเลือกใหม่สำหรับการเรียนการสอนมากขึ้น

3. วิธีการดำเนินการวิจัย

การวิจัยนี้นำเอาเทคโนโลยีเข้าร่วมกับการจัดการเรียนการสอนผ่านเครือข่ายอินเทอร์เน็ต โดยใช้รูปแบบการสอนในมหาวิทยาลัยราชภัฏนครราชสีมาเป็นต้นแบบระบบ ดังภาพที่ 1



ภาพที่ 1: กระบวนการทำงานโดยรวมของระบบ

โดยขอบเขตของงานวิจัยแบ่งออกเป็น 3 ส่วน คือ ขอบเขตด้านผู้ใช้งาน คือ อาจารย์ระดับปริญญาโทสาขาคอมพิวเตอร์ นักศึกษาโปรแกรมวิชาคอมพิวเตอร์ศึกษาและวิทยาการสารสนเทศ มหาวิทยาลัยราชภัฏนครราชสีมา ขอบเขตด้านผลลัพธ์ คือ การพัฒนาระบบจัดการเรียนการสอนแบบออนไลน์สำหรับอาจารย์ นักศึกษามหาวิทยาลัยราชภัฏนครราชสีมาและความพึงพอใจในการใช้งานระบบ ขอบเขตด้านเครื่องมือที่ใช้ คือ แบบสัมภาษณ์และแบบสอบถาม

วิธีการดำเนินงานวิจัยแบ่งเป็น 4 ขั้นตอน คือ 1. ศึกษาและวิเคราะห์ระบบเดิมที่มีอยู่ 2. การวิเคราะห์และออกแบบระบบใหม่ 3. พัฒนาระบบการจัดการเรียนการสอน และ 4. ทดลองใช้และประเมินผล จากการสัมภาษณ์และแบบสอบถาม

3.1 ขั้นที่หนึ่งวิเคราะห์ระบบเดิม

การวิเคราะห์ระบบเดิมประกอบด้วย 3 ส่วน คือ การศึกษาจากกรณีตัวอย่าง จากการสัมภาษณ์และจากแบบสอบถาม ดังนี้

3.1.1 กรณีศึกษา ได้ทำการศึกษาระบบการจัดการเรียนการสอนออนไลน์ของมหาวิทยาลัยราชภัฏนครราชสีมาแบบเดิมที่มีอยู่ โดยในระบบประกอบไปด้วยการให้บริการแก่อาจารย์และ

นักศึกษา ได้แก่ การสร้างรายวิชา การจัดการเรียนการสอน การนำเข้าคำถาม คะแนน เว็บบอร์ด เป็นต้น ดังภาพที่ 2



ภาพที่ 2: ระบบการจัดการเรียนการสอนออนไลน์แบบเดิม

โดยมีการกำหนดขนาดของกลุ่มตัวอย่าง คือ ประชากรที่ถูกเลือกด้วยวิธีการต่างๆ เพื่อนำมาใช้ในทดลองและจะมีผลต่อการประเมินผลงานวิจัย ซึ่งมีผลกระทบต่อความคลาดเคลื่อนของผลการวิจัย ความสำคัญของการหาขนาดของกลุ่มตัวอย่างเพื่อใช้วัดประสิทธิภาพของงานวิจัย Taro Yamane [9] ได้นำเสนอสูตรการคำนวณขนาดอย่างสัดส่วน 1 กลุ่มที่มีจำนวนประชากรแบบจำกัดหรือรู้จำนวนประชากรที่แน่นอน (Finite Population) และมีการกำหนดค่าความคลาดเคลื่อนของกลุ่มตัวอย่างระหว่างค่าที่แท้จริงกับค่าที่ประมาณที่ผู้วิจัยสามารถยอมรับได้มาใช้คำนวณให้ได้ผลลัพธ์ที่มีความแม่นยำมากที่สุด ดังสมการที่ 1

$$n = \frac{N}{1 + NE^2} \tag{1}$$

เมื่อกำหนดให้ n คือ ขนาดของกลุ่มตัวอย่าง, N คือ จำนวนประชากร และ E คือ ความคลาดเคลื่อนที่ยอมรับให้เกิดขึ้นในรูปของสัดส่วน [9] และได้กำหนดระดับความพึงพอใจต่อระบบการจัดการเรียนการสอนเป็นมาตรฐานอันดับ 5 ระดับ ดังตารางที่ 1

ตารางที่ 1: เกณฑ์การให้คะแนนของความพึงพอใจ

ระดับคะแนน (เชิงปริมาณ)	ระดับความพึงพอใจต่อระบบ (เชิงคุณภาพ)
5	ดีมาก
4	ดี
3	พอใช้
2	ปรับปรุง
1	ไม่เหมาะสม

3.1.2 การสัมภาษณ์ ได้สัมภาษณ์อาจารย์ที่ใช้งานระบบการจัดการเรียนการสอนเดิม จำนวน 5 ท่าน โดยแบ่งข้อสัมภาษณ์เป็น 2 ส่วน คือ คำถามแบบปลายปิด เพื่อให้ผู้ถูกสัมภาษณ์ตอบตามความสนใจของผู้วิจัย และคำถามแบบปลายเปิด เพื่อให้ผู้ถูกสัมภาษณ์เสนอแนวความคิดหรือความต้องการต่อระบบ ซึ่งสรุปปัญหาของระบบเดิมได้คือ รูปแบบและเงื่อนไขในการใช้งานมีความซับซ้อนเกินความต้องการ

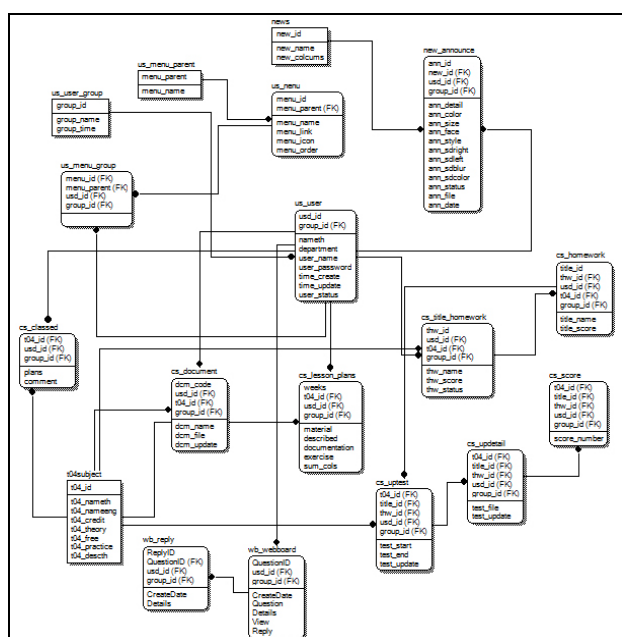
3.1.3 แบบสอบถาม ใช้กับนักศึกษา แบบเจาะจง จากจำนวนนักศึกษาทั้งสิ้น 420 คน ซึ่งให้ความคลาดเคลื่อนของกลุ่มตัวอย่างไม่เกินร้อยละ 5 และจากสมการที่ 1 ขนาดของกลุ่มตัวอย่างจึงเท่ากับ 204 คน มาทำแบบสอบถามเกี่ยวกับระบบเดิม สรุปผลได้คือระบบยุ่งยากต่อการใช้งาน การแสดงผลไม่สวยงามและจำนวนเมนูมีมากเกินไปจนความจำเป็น

3.2 ขั้นตอนที่สองการวิเคราะห์และออกแบบระบบใหม่

เป็นขั้นตอนการวิเคราะห์และออกแบบระบบใหม่ ซึ่งประกอบไปด้วยการทำงานดังต่อไปนี้

3.2.1 ศึกษาและวิเคราะห์รูปแบบการสอน การจัดเก็บคะแนน การใช้งานระบบแบ่งออกเป็น 3 ส่วน คือ ผู้ดูแลระบบ อาจารย์และนักศึกษา

3.2.2 ออกแบบระบบการทำงานภายในด้วย Data Flow diagram ออกแบบหน้าจอโดยใช้ Storyboard และออกแบบฐานข้อมูลด้วย E-R Diagram โดยใช้ MySQL เป็นฐานข้อมูลของระบบ โครงสร้างภายในฐานข้อมูลแสดงได้ดังภาพที่ 3



ภาพที่ 3: ER-Diagram ระบบการจัดการเรียนการสอนออนไลน์

3.3 ขั้นตอนที่สามการพัฒนาโปรแกรม

เป็นการพัฒนาระบบจัดการเรียนการสอนเป็นเว็บแอปพลิเคชัน โดยใช้ภาษา PHP , SQL ในการเชื่อมต่อกับฐานข้อมูล โดยระบบประกอบไปด้วย 6 ระบบ คือ ระบบข้อรายวิชา ระบบข้อมูลรายวิชา ระบบควบคุมผู้ใช้ ระบบจัดเก็บคะแนน ระบบแก้ไขข้อมูล และ กระบวนการตอบ

3.4 ขั้นตอนที่สี่การทดสอบและประเมินผล

การทดสอบระบบเพื่อหาข้อบกพร่องและควบคุมคุณภาพของระบบ การประเมินความพอใจของระบบใช้การประเมินความพึงพอใจด้วยมาตรอันดับ 5 อันดับ และแปลความหมายของค่าเฉลี่ยเป็นแบบช่วง 5 ช่วงคะแนน [10] ดังตารางที่ 2

ตารางที่ 2: เกณฑ์การแปลผลการประเมินความพึงพอใจ

เกณฑ์ค่าเฉลี่ย	การแปลความหมายของระดับคะแนน
4.51–5.00	ความพึงพอใจอยู่ในเกณฑ์ดีมาก
3.51–4.50	ความพึงพอใจอยู่ในเกณฑ์ดี
2.51–3.50	ความพึงพอใจอยู่ในเกณฑ์พอใช้
1.51–2.50	ความพึงพอใจอยู่ในเกณฑ์ปรับปรุง
1.00–1.50	ความพึงพอใจอยู่ในเกณฑ์ไม่เหมาะสม

ซึ่งสถิติที่ใช้ในการวิเคราะห์ข้อมูลความพึงพอใจต่อระบบมี 2 สมการ คือ ค่าเฉลี่ยเลขคณิต (Mean) ดังสมการที่ 2 และส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) ดังสมการที่ 3

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \tag{2}$$

$$S.D. = \sqrt{\frac{n \sum fx^2 - (\sum x)^2}{n(n-1)}} \tag{3}$$

การประเมินความพอใจต่อการใช้ระบบแบ่งเป็น 2 ส่วน คือ

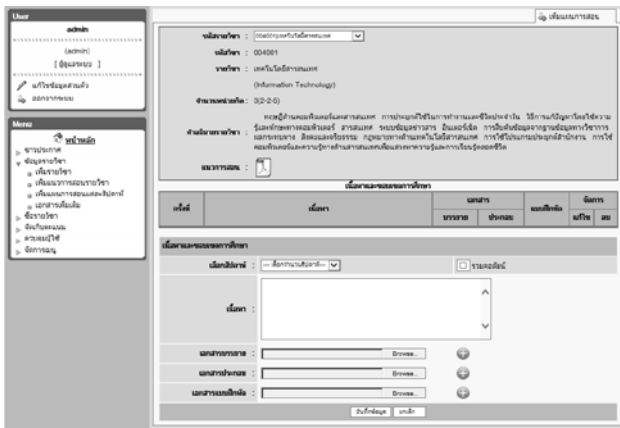
3.4.1 ผลการสัมภาษณ์ของอาจารย์ โดยนาระบบให้อาจารย์ทดลองใช้และประเมินผลเป็นภาพรวมของการใช้งานผลได้จากการสัมภาษณ์จากคำถามแบบปลายเปิดและปลายปิด

3.4.2 ผลของแบบสอบถามจากนักศึกษา ซึ่งแบบสอบถามแบ่งออกเป็น 2 ส่วน คือ ส่วนขององค์ประกอบภายในระบบและส่วนของการใช้งานระบบ

4. ผลการดำเนินการวิจัย

4.1 ส่วนของระบบ

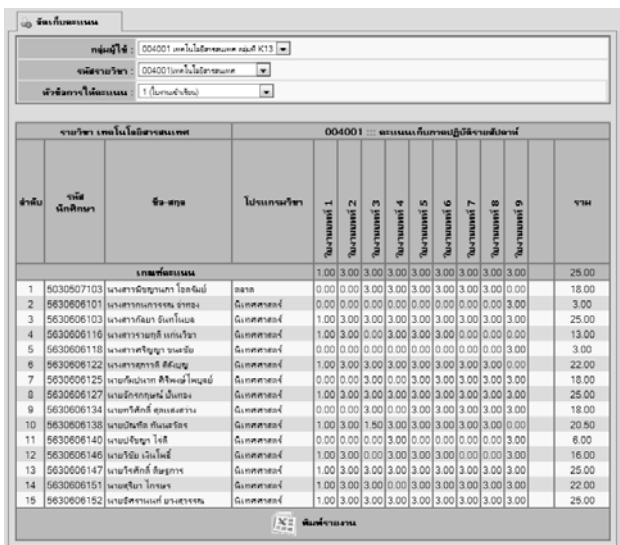
ระบบการจัดการเรียนการสอนออนไลน์สำหรับอาจารย์และนักศึกษามหาวิทยาลัยราชภัฏนครราชสีมา ได้ผลดังนี้



ภาพที่ 4: ส่วนของการจัดการเนื้อหาในรายวิชาโดยผู้ดูแลระบบ



ภาพที่ 5: ส่วนของการจัดการคะแนนของอาจารย์



ภาพที่ 6: ส่วนของการแสดงผลคะแนนเก็บของนักศึกษา

4.2 ส่วนของผลการทดลองใช้

ผลการประเมินแบ่งออกเป็น 2 ส่วน คือ

4.2.1 ความพึงพอใจของอาจารย์ มีความพึงพอใจต่อระบบใหม่อยู่ในระดับดี ค่าเฉลี่ยเท่ากับ 4.01 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.83 เพราะระบบสามารถใช้งานได้ง่ายและตรงกับความต้องการ แต่ยังไม่ครอบคลุมการใช้งานทั้งหมดในการจัดกิจกรรมการเรียนการสอน

4.2.2 ความพึงพอใจของนักศึกษา แสดงได้ดังตารางที่ 3

ตารางที่ 3: เปรียบเทียบผลความพึงพอใจในระบบของนักศึกษา

รายละเอียด	ระบบเดิม	ระบบใหม่	S.D.
ค่าเฉลี่ยความพอใจ			
ส่วนขององค์ประกอบระบบ			
ความชัดเจนของหัวข้อใช้งาน	4.42	4.59	0.58
จำนวนหัวข้อการใช้งานพอเหมาะ	4.29	4.22	0.65
การจัดหน้าจอเป็นระเบียบสวยงาม	4.19	4.24	0.80
ส่วนช่วยเหลือหรือคำอธิบายเพิ่ม	4.18	4.13	0.76
ส่วนของการใช้งาน			
ตรงกับความต้องการของผู้ใช้	4.37	4.41	0.72
รูปแบบใช้งานและง่ายต่อการเรียนรู้	4.25	4.37	0.71
ประโยชน์ที่ได้รับจากการใช้งาน	4.49	4.34	0.69
ความพึงพอใจโดยรวมต่อระบบ	4.37	4.38	0.76
ค่าเฉลี่ยโดยรวมของระบบ	4.32	4.34	0.71

จากตารางที่ 2 สรุปได้ว่า ระบบจัดการเรียนการสอนแบบออนไลน์สำหรับอาจารย์และนักศึกษามหาวิทยาลัยราชภัฏนครราชสีมาที่ผู้วิจัยพัฒนาขึ้น นักศึกษามีความพึงพอใจโดยรวมต่อระบบอยู่ในระดับดี โดยมีค่าเฉลี่ยเท่ากับ 4.34

5. สรุป

จากการปัญหาการจัดการเรียนการสอนของมหาวิทยาลัยราชภัฏนครราชสีมาที่มีรูปแบบที่หลากหลาย ทำให้การเก็บคะแนนหรือการจัดการเรียนการสอนทำได้ยาก เนื่องจากนักศึกษามีความสามารถต่างกันและผู้สอนมีวิธีการสอนต่างกัน อีกทั้งระบบการจัดการเรียนการสอนเดิมมีความซับซ้อนใช้งานได้ยากและไม่ตรงกับความต้องการของบุคลากรในมหาวิทยาลัย

ราชภัฏนครราชสีมา ดังนั้นในงานวิจัยนี้ได้นำเสนอการพัฒนา
ระบบจัดการเรียนการสอนแบบออนไลน์สำหรับอาจารย์และ
นักศึกษามหาวิทยาลัยราชภัฏนครราชสีมา โดยการพัฒนาเป็น
เว็บแอปพลิเคชันผ่านระบบเครือข่ายอินเทอร์เน็ต ช่วยแก้ปัญหา
การเก็บคะแนน การจัดกิจกรรมการเรียนการสอนให้ตรงกับ
ความต้องการของผู้สอน และการใช้งานที่ง่ายของนักศึกษาหรือ
อาจารย์ ซึ่งผลลัพธ์ของงานวิจัยนี้มีค่าเฉลี่ยความพึงพอใจของ
อาจารย์ต่อระบบใหม่เท่ากับ 4.01 และส่วนเบี่ยงเบนมาตรฐาน
เท่ากับ 0.83 ซึ่งอยู่ในเกณฑ์ดีและค่าเฉลี่ยความพอใจของ
นักศึกษาต่อระบบใหม่เท่ากับ 4.34 และส่วนเบี่ยงเบนมาตรฐาน
เท่ากับ 0.71 ซึ่งอยู่ในเกณฑ์ดี

เอกสารอ้างอิง

- [1] ทศนีย์ รอดมันคง และ ปณิดา วรรณพิรุณ, “การพัฒนารูปแบบ
กิจกรรมการเรียนการสอนโดยใช้คอมพิวเตอร์สนับสนุนการเรียนรู้
แบบร่วมมือตามแนวการจัดการเรียนรู้แบบลดภาระทางปัญญาเพื่อ
พัฒนากระบวนการแลกเปลี่ยนเรียนรู้”, *การประชุมวิชาการ
ระดับชาติด้านอิเล็กทรอนิกส์ นวัตกรรมและการเรียนรู้ออนไลน์ประชาคม
อาเซียน: นโยบายและกระบวนการ (Integrating ASEAN
Online learning: Policy and Process)*, สิงหาคม 2555, หน้า
243-250.
- [2] เสมอกาญจน์ ไสภณศิริภรณ์ และ ปราวีณา สุวรรณณัฐโชติ,
“มุมมองของผู้สนับสนุนการเปลี่ยนแปลงในโปรแกรมอิเล็กทรอนิกส์
เพื่อการส่งเสริมการนำการเรียนการสอนแบบอิเล็กทรอนิกส์ไปใช้ใน
ระดับอุดมศึกษา (Perspective of Change Facilitators in e-
Learning Program for Promoting e-Learning
Implementation in Higher Education)”, *การประชุมวิชาการ
ระดับชาติด้านอิเล็กทรอนิกส์ นวัตกรรมและการเรียนรู้ออนไลน์ประชาคม
อาเซียน: นโยบายและกระบวนการ (Integrating ASEAN
Online learning: Policy and Process)*, สิงหาคม 2555, หน้า
73-80.
- [3] สุชาดา สามสวัสดิ์, *E-learning กับการศึกษาไทย (E-learning
with Thai Education)*, ภาควิชาเทคโนโลยีและการสื่อสาร
การศึกษามหาวิทยาลัยนครราชสีมา, 2555.
- [4] วิเชียร ขุนเพชร, บทความ เรื่อง E-learning คืออะไร, สำนักวิทย
บริการ, มหาวิทยาลัยทักษิณ วิทยาเขตพัทลุง.
- [5] ระบบการศึกษาทางไกล E-Learning. [ออนไลน์]. วันที่เข้าถึง
ข้อมูล 10 มีนาคม 2556. เข้าถึงได้จาก :
[http://www.netthailand.com/home/articles.php?art_id=12
&start=1](http://www.netthailand.com/home/articles.php?art_id=12&start=1).
- [6] การพัฒนาระบบจัดการเรียนการสอนรูปแบบ-Learning ของ
สพฐ. (OBEC LMS) ตอนที่3 MMS และLMS. [ออนไลน์].
วันที่เข้าถึงข้อมูล 7 มีนาคม 2556. เข้าถึงได้จาก :
[http://anekrati.wordpress.com/2010/09/28/การพัฒนาระบบ
จัดการเรียน-4/](http://anekrati.wordpress.com/2010/09/28/การพัฒนาระบบจัดการเรียน-4/).
- [7] ฉมนน จีรังสุวรรณ และ ธนยศ สิริโชค, “รูปแบบการเรียนการ
สอนแบบร่วมมือด้วยระบบการจัดการกิจกรรมการเรียนรู้ (Model of
Collaborative Learning Using Learning Activity
Management System)”, *การประชุมวิชาการระดับชาติด้านอิเล็ก
ทรอนิกส์ นวัตกรรมและการเรียนรู้ออนไลน์ประชาคมอาเซียน: นโยบาย
และกระบวนการ (Integrating ASEAN Online learning:
Policy and Process)*, สิงหาคม 2555, หน้า 123-130.
- [8] พิชิต ศรีวิทย์รัตน์ และ วรรณภา ศรีวิทย์รัตน์, “นวัตกรรมจัดการ
เรียนการสอนระบบอิเล็กทรอนิกส์แบบจุดเดียวเบ็ดเสร็จ (One Stop e-
Learning Management Innovation)”, *การประชุมวิชาการ
ระดับชาติด้านอิเล็กทรอนิกส์ นวัตกรรมและการเรียนรู้ออนไลน์ประชาคม
อาเซียน: นโยบายและกระบวนการ (Integrating ASEAN
Online learning: Policy and Process)*, สิงหาคม 2555, หน้า
110-113.
- [9] Taro. Yamane, 2nd Ed. *Statistics, An Introductory
Analysis*, Harper and Row: New York, 1967.
- [10] ทศนวรรณ แก้วใส และ สุพจน์ นิตย์สุวัฒน์, “ระบบแนะนำ
ภาพยนตร์ด้วยเทคนิคตัวกรองเชิงร่วมมือร่วมกับวิธีเคมีน (Movies
Recommender System using Collaborative Filtering and
K-Means)”, *The 5th National Conference on Computing
and Information Technology (NCCIT 2009)*, พฤษภาคม
2552, หน้า 502-507.

การจัดตารางสอบด้วยวิธีการบรรจุงานกรณีศึกษาคณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น

The Examination Timetable Arrangement by Bin Packing: A Case Study of Faculty of Science, Khon Kaen University

จรรยาลักษณ์ สุพร (Jaroonluk Suporn)¹, สิริภัทร เชี่ยวชาญวัฒนา (Sirapat Chiewchanwattana)²

และคำรณ สุนัตติ (Khamron Sunat)³

ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น

¹jaroonluk@kku.ac.th, ²sunkra@kku.ac.th, ³khamron_sunat@yahoo.com

บทคัดย่อ

บทความนี้นำเสนอวิธีการแก้ปัญหาการจัดตารางสอบด้วยเทคนิคการบรรจุงานที่มีการใช้ทรัพยากรที่มีประสิทธิภาพ วิธีการที่นำเสนอแบ่งเป็น 2 ส่วน ส่วนแรกคือ การเลือกอาคารที่ใช้ในการจัดสอบโดยพิจารณาจากจำนวนที่นั่งสอบรวมทั้งหมดของอาคารนั้น ส่วนที่สองคือการจัดวิชาลงในห้องสอบ เป้าหมายเพื่อให้ค่าใช้จ่ายน้อยที่สุด โดยคำนึงถึงค่าตอบแทนกรรมการคุมสอบ และค่าไฟฟ้าที่ใช้จัดสอบ ซึ่งปัญหาการจัดตารางสอบเดิมนี้มีความซับซ้อนและต้องพิจารณาถึงเงื่อนไขและปัจจัยหลายประการ เช่น จำนวนรายวิชา จำนวนนักศึกษา จำนวนที่นั่งสอบ ค่าใช้จ่ายที่จะเกิดขึ้นในแต่ละห้องสอบและข้อบังคับต่างๆ ที่ใช้ในการจัดตารางสอบด้วย โดยชุดข้อมูลที่นำมาใช้ในการทดลอง ใช้ข้อมูลรายวิชา, เวลาในการจัดสอบและข้อมูลห้องสอบ ในคณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น 2 ภาควิชาการศึกษ ผลเปรียบเทียบระหว่างวิธีที่นำเสนอกับวิธีการจัดแบบเดิมที่ต้องใช้ผู้เชี่ยวชาญในการจัดห้องสอบ ผลการทดลองพบว่าวิธีการที่นำเสนอเพื่อใช้จัดตารางสอบนั้น สามารถลดเวลาและค่าใช้จ่ายที่ใช้ในการจัดสอบ

คำสำคัญ: ตารางสอบ วิธีการบรรจุงาน

Abstract

This paper presents the solution of examination timetable arrangement based on Bin Packing technique with efficiency using of resources. The proposed method is separated into 2 parts. The first part, the suitable building is determined by considering the total number of seats in the building. The second part is assigning the subjects to rooms. The proposed method aims to reduce

the expense of invigilation and electricity bill. The original timetable arrangement problem is quite complicated. There are many conditions and factors, such as the total numbers of subject, the total numbers of room, the total number of seats, room cost, and the other regulations, in determining the suitable rooms for subjects. The data set used in experiment is the details of subject, room, the schedule time of examination arrangement of 2 semester of Faculty of Science, Khon Kaen University. The results shows that the proposed method outperforms the existing method in terms of cost savings and time reduction.

Keywords: Examination Timetable, Bin Packing.

1. บทนำ

การจัดตารางสอบถือเป็นปัญหาที่มีความสำคัญอย่างมากในกระบวนการจัดการศึกษา แม้จะเป็นงานที่ต้องทำเป็นประจำในทุกภาคการศึกษา แต่ด้วยปัจจัยที่แตกต่างกันในการจัดสอบ เช่น จำนวนรายวิชา จำนวนห้องที่ใช้ จำนวนนักศึกษาที่มีสิทธิสอบในแต่ละรายวิชา ทำให้การจัดตารางสอบยังคงเป็นปัญหาที่มีผู้ให้ความสนใจหาวิธีแก้ปัญหา ซึ่งบางหน่วยงานที่ใช้บุคลากรจัดตารางสอบนั้นถึงแม้ว่าจะมีความเชี่ยวชาญ แต่ก็ยังคงเป็นเรื่องที่ยุ่งยากและซับซ้อน ในปัจจุบันจึงได้มีการนำเอาหลักการและแนวคิดทางเทคโนโลยีสารสนเทศต่างๆ มาประยุกต์ใช้ในการจัด

ปัญหาในการจัดตารางสอบที่มีความซ้ำซ้อนหรือชนกัน กล่าวคือ ใน 1 ช่วงเวลาสอบใดๆ มีวิชาที่จัดสอบในห้องสอบเดียวกันมากกว่า 1 รายวิชาซึ่งถือว่าเป็นเงื่อนไขหลักของการจัด

ตารางสอบ เพราะที่นั่งสอบแต่ละที่ไม่สามารถสอบได้พร้อมกัน 2 รายวิชา จะเห็นได้จากงานวิจัยหลายๆ งาน เช่นวราวุธ อินทวงษ์ [1] ได้ทดลองการจัดตารางสอบหัวข้อวิจัยโดยใช้เทคนิควิธีไบเซ็คติงเคมินส์, วนาลี แก้วพะนอ, สุทามาศ มนตรีบริรักษ์ [2] ที่ได้ทดลองการจัดตารางสอบหัวข้อวิจัยเพื่อให้คาบเวลาสอบเสร็จเฉลี่ยน้อยสุด, อรุณี ปุกแก้ว [3] ได้ทดลองจัดตารางสอนตารางสอบขึ้นต้นก่อนนำไปจัดตารางร่วมกับคณะวิชาอื่นๆ, มนต์สวรรค์ ชันโท [4] ได้ทดลองจัดตารางสอนตารางสอบ เพื่อแก้ปัญหาการจัดผู้สอบไม่ให้เกิน 2 คน, นพพล บุณย์กุศล [5] ได้ทดลองการจัดตารางสอบประจำภาคเรียนโดยใช้อัลกอริทึมแบบไฮบริด, ประกิจพันธ์ พัฒน-หิรัญธำรง [6] ได้ทดลองจัดตารางสอบเพื่อลดเวลาในการดำเนินการจัดตารางสอบ และสุรเดช ศรีวิทยาภรณ์ [7] ได้ทดลองจัดตารางสอบหัวข้อวิจัยโดยอาศัยเทคนิค Genetic Algorithms จากงานวิจัยที่ได้กล่าวมาจะเห็นได้ว่าส่วนใหญ่การจัดจะมีวัตถุประสงค์แตกต่างกันไป และเน้นในด้านการประหยัดทรัพยากรเป็นส่วนใหญ่

เทคนิคการบรรจุงาน (Bin Packing) เป็นเทคนิคที่นิยมนำไปใช้แก้ปัญหาหลายอย่าง เช่น การจัดตารางงาน การจัดแบ่งทรัพยากร การแยกส่วนวัสดุ เป็นต้น แต่จากปัญหาการจัดตารางสอบเป็นปัญหาที่ไม่สามารถหาคำตอบได้ในเวลาโพลิโนเมียลจัดเป็นปัญหาเอ็นพีแบบยาก NP-HARD [8] ซึ่งทำให้การแก้ปัญหาการจัดตารางสอบทำได้ไม่ดีหากคำนึงถึงค่าใช้จ่ายที่จะเกิดขึ้น และการใช้ทรัพยากรที่ยังไม่คุ้มค่าเท่าที่ควร ในบทความนี้จึงได้เสนอขั้นตอนวิธีการแก้ปัญหาการจัดตารางสอบด้วยวิธีการบรรจุงานแบบใหม่ โดยศึกษาผลกระทบด้านการใช้ทรัพยากรการจัดสอบให้คุ้มค่า ข้อดีของ Bin Packing ได้ที่นั่งเหลือน้อยเมื่อใช้ในการจัดสอบ แต่ปัญหาคือการจัดสอบของคณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น มีเงื่อนไขที่ซับซ้อนที่ไม่สามารถใช้เทคนิคเดียวในการจัดสอบได้ ดังนั้นในบทความนี้จึงได้เสนอการจัดตารางสอบแบบบรรจุงานแบบผสมผสานร่วมกับเทคนิค First-Fit และ Best-Fit

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การจัดตารางงาน

การจัดตารางงาน (Task Scheduling) สามารถแบ่งออกเป็นสองวิธี 1) การจัดตารางการทำงานแบบแบ่งส่วน (Partitioning

Scheduling) เมื่องานถูกจัดสรรให้กับโพรเซสเซอร์ตัวใดแล้วงานนั้นจะไม่สามารถย้ายไปประมวลผลที่โพรเซสเซอร์อื่นได้ 2) การจัดตารางการทำงานแบบครอบคลุม (Global Scheduling) [9] วิธีนี้งานแต่ละงานสามารถย้ายไปประมวลผลได้หลายโพรเซสเซอร์จนกว่างานนั้นจะประมวลผลเสร็จสมบูรณ์ ข้อดีและข้อเสียของการจัดตารางงานแบบแบ่งส่วนและแบบครอบคลุม เช่นการจัดตารางการทำงานแบบแบ่งส่วนมีความซับซ้อนและโอเวอร์เฮด (Overhead) น้อยกว่าการจัดตารางการทำงานแบบครอบคลุม อย่างไรก็ตามการจัดตารางแบบครอบคลุมสามารถย้ายงานไปประมวลผลที่โพรเซสเซอร์ต่างๆ ได้ทำให้มีประสิทธิภาพในการใช้โพรเซสเซอร์ของระบบ (System Utilization) ดีกว่าแบบแบ่งส่วน โดยอาศัยแผนการจัดบรรจุแบบครอบคลุม (Global Bin Packing) ซึ่งผลจากการจำลองการทำงานนั้นแสดงให้เห็นถึงประสิทธิภาพที่ดีขึ้นของการจัดตารางการทำงาน

2.2 การจัดวางบรรจุภัณฑ์ในอุปกรณ์บรรจุ (Bin-Packing)

การบรรจุวัตถุสี่เหลี่ยมลงในพื้นที่สี่เหลี่ยมที่มีปริมาตรไม่จำกัด [10] โดยมีการกำหนดปริมาตรของวัตถุ จำนวนในการบรรจุ และปริมาตรของพื้นที่บรรจุและจัดวาง ซึ่งมีวัตถุประสงค์ในการจัดวางโดยใช้จำนวนพื้นที่สี่เหลี่ยมในการบรรจุให้น้อยที่สุด ซึ่งได้มีการกำหนดสมการวัตถุประสงค์ ดังนี้

X_{ij} เป็น 1 ถ้าพื้นที่สี่เหลี่ยม i ถูกเลือกให้บรรจุวัตถุ j เป็น 0 ถ้าไม่ถูกเลือก, Y_i เป็น 1 ถ้าพื้นที่สี่เหลี่ยม i ถูกเลือกให้บรรจุวัตถุ เป็น 0 ห้ามบรรจุวัตถุ, t_j คือ ปริมาตรวัตถุ j , p_i คือ ปริมาตรพื้นที่สี่เหลี่ยม

$$\text{Minimize } \sum_{i=1}^n Y_i \quad (1)$$

$$\text{Subject to } \sum_{j=1}^m t_j X_{ij} \leq p_i \quad i=1, \dots, n \quad (2)$$

$$\sum_{i=1}^n X_{ij} = 1 \quad j=1, \dots, m \quad (3)$$

$$X_{ij} \leq Y_i \quad \text{for all } i, j \quad (4)$$

$$X_{ij} = \{0,1\}, Y_i = \{0,1\} \quad (5)$$

2.3 การจัดสรรเนื้อที่แบบพหุ (Multiple-Partition Allocation)

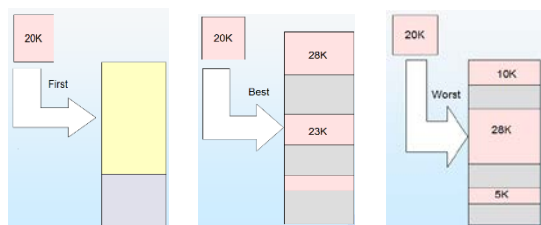
ในระบบปฏิบัติการจะมีตารางที่เก็บข้อมูลว่าพื้นที่ในหน่วยความจำหลักส่วนใดถูกครอบครองและส่วนใดยังว่างอยู่ [11] เริ่มต้นพื้นที่ในหน่วยความจำจะว่างทั้งหมด เมื่อมีกระบวนการเข้ามาในระบบ ระบบจะจัดสรรพื้นที่ที่ใหญ่พอ

สำหรับกระบวนการนั้น และจัดแบ่งพื้นที่ให้กระบวนการเพียงเท่าที่ที่ต้องการ ส่วนพื้นที่ที่เหลือจะเก็บไว้ใช้ในการจัดสรรครั้งต่อไป ซึ่งวิธีการจัดหาพื้นที่ว่างมีหลายวิธี จากภาพที่ 1 แต่ละวิธีพยายามให้การจัดสรรพื้นที่มีประโยชน์สูงสุดวิธีที่นิยมได้แก่

First-Fit คือ การเลือกพื้นที่ที่มีขนาดใหญ่กว่าหรือเท่ากับพื้นที่ที่ต้องการ โดยเริ่มค้นหาตั้งแต่จุดเริ่มต้นของกลุ่มพื้นที่ว่างหรือค้นหาจากจุดที่หยุดลงจากการค้นหาครั้งก่อน เพื่อให้การเลือกห้องสอบอยู่ในอาคารเดียวกัน ทำงานได้เร็วข้อมูลจะกระจุกอยู่ด้านบน และอาจเกิด Fragment ที่มีขนาดใหญ่ได้

Best-Fit คือ การเลือกพื้นที่ขนาดใกล้เคียงกับขนาดพื้นที่ที่ต้องการมากที่สุด การค้นหาพื้นที่ว่างแบบนี้จะต้องตรวจสอบพื้นที่ว่างทั้งหมดในระบบ ถ้าได้มีการเรียงพื้นที่ว่างไว้ตามลำดับขนาดแล้ว การค้นหาจะเร็วขึ้น วิธีนี้จะทำให้พื้นที่ว่างที่เกิดขึ้นใหม่มีขนาดเล็กที่สุด ระบบทำงานช้า เพราะคำนวณมาก และเกิด Fragment ขนาดเล็ก แต่มีจำนวนมาก

Worst-Fit คือ การเลือกพื้นที่ที่มีขนาดใหญ่กว่าขนาดพื้นที่ที่ต้องการมากที่สุด การค้นหาจะต้องตรวจสอบพื้นที่ว่างทั้งหมดในระบบ นอกจากจะมีการเรียงไว้แล้ว ซึ่งการเลือกวิธีนี้ทำให้เกิดพื้นที่ว่างใหม่ที่มีขนาดใหญ่ที่สุด อาจนำไปใช้ต่อไปได้มากกว่าวิธี Best-Fit ซึ่งทำให้พื้นที่ว่างที่เกิดขึ้นใหม่มีขนาดเล็กที่สุด ระบบทำงานช้า เพราะคำนวณมากโอกาสเกิด Fragment น้อยลง



ภาพที่ 1: แสดงวิธีวิธีการจัดหาพื้นที่ในหน่วยความจำหลัก

3. วิธีดำเนินการวิจัย

บทความนี้ได้นำเสนอวิธีการจัดตารางสอบด้วยเทคนิคการบรรจุนานที่ได้ผสมผสานกันระหว่างขั้นตอนวิธีการบรรจุนานทั้ง First-Fit และ Best-Fit โดยใช้ข้อมูลการจัดตารางสอบของคณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น เป็นข้อมูลในการทดสอบจริง มีรายละเอียดในการดำเนินการวิจัยดังนี้

3.1 การเตรียมข้อมูล

ข้อมูลที่ใช้ในการจัดตารางสอบ ได้แก่ข้อมูลวิชาสอบนักศึกษา (ตารางที่ 1) 2 ภาคการศึกษา คือ กลางภาค/ปลายภาค

ภาคการศึกษาที่ 1 ปีการศึกษา 2555, กลางภาค/ปลายภาค ภาคการศึกษาที่ 2 ปีการศึกษา 2555 และ ข้อมูลห้องเรียนคณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น (ตารางที่ 2)

ตารางที่ 1: ตัวอย่างข้อมูลรายวิชา

ลำดับ	วันที่	เวลา	วิชา	นศ.ปกติ	นศ.พิเศษ
1	WBA	13:00	311101	214	0
2	WBA	13:00	311111	354	106
3	23 ก.ค. 2555	13:00	311244	49	0
4	23 ก.ค. 2555	13:00	314201	51	0
5	23 ก.ค. 2555	08:30	314123	211	0

ข้อมูลที่นำมาใช้ในการทดลอง ใช้ข้อมูลจาก ม.ข.30 ตารางที่ 1 และข้อมูลห้องสอบคณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น เป็นข้อมูลที่ได้มาจากหน่วยอาคารสถานที่และยานพาหนะตามตารางที่ 2

ตารางที่ 2: ตัวอย่างข้อมูลห้องสอบ

ที่	อาคาร	ห้อง	ที่นั่งสอบ		พลังงานไฟฟ้า 3 ชั่วโมง /KW	
			A	B	แอร์	แสงสว่าง
1	กล่ม(SC.01)	SC1103	39	38	66.24	21.05
2	กล่ม(SC.01)	SC1101	39	38	66.24	21.25
3	กล่ม(SC.01)	SC1102	39	38	66.24	21.25
4	ฟิลิส(SC.02)	SC2101	45	0	70.06	2.31
5	ฟิลิส(SC.02)	SC2201	45	0	49.27	1.5015

3.2 การวิเคราะห์ข้อมูล

สำหรับขั้นตอนวิธีการจัดตารางสอบของคณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น จะพิจารณาวันสอบที่กำหนดโดย ม.ข.30 เป็นหลัก และต้องเป็นวิชาที่เวลาสอบไม่เป็น WBA (ผู้สอนหรือภาควิชาจัดสอบเอง) เงื่อนไขในการจัดตารางสอบจะแบ่งเป็น 2 เงื่อนไข คือเงื่อนไขหลักและเงื่อนไขรอง ดังนี้

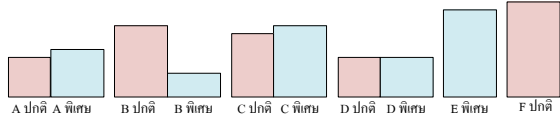
เงื่อนไขหลัก : 1. วันสอบและวิชาที่สอบกำหนดตาม ม.ข.30 ไม่สามารถแก้ไขหรือเพิ่มลบจำนวนได้ 2.ผู้คุมสอบ จัดได้ดังนี้ (1) จำนวน นศ. ไม่เกิน 129 คน: ผู้คุมสอบ 2 ราย (2) จำนวน นศ. 130-199 คน: ผู้คุมสอบ 3 ราย (3) จำนวน นศ. ตั้งแต่ 200 คนขึ้นไป: ผู้คุมสอบ 4 ราย

เงื่อนไขรอง : 1.จัดสอบวิชาเดียวกัน ไม่สลับแถว 2.นักศึกษาภาคปกติ – โครงการพิเศษ จัดสอบในห้องเดียวกัน แต่คิดค่าคุมสอบให้กรรมการคุมสอบ

การจัดตารางสอบด้วยวิธีการบรรจุนานนี้ไม่พิจารณาวิชาที่สอบชนกันและมีขั้นตอนการจัดตารางสอบดังนี้

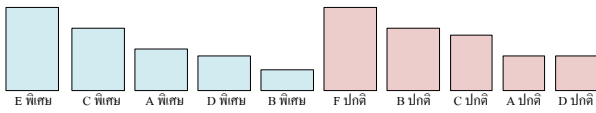
1. หาผลรวมจำนวนที่นั่งสอบแต่ละอาคาร และจำนวนนักศึกษาทั้งหมด

- เลือกอาคารในการจัดสอบ โดยมีจำนวนที่นั่งสอบเพียงพอกับจำนวนนักศึกษา (First-Fit)
- เตรียมข้อมูลรายวิชาที่ใช้ในการทดสอบ คือข้อมูลการสอบแต่ละวิชาจาก ม.ข.30



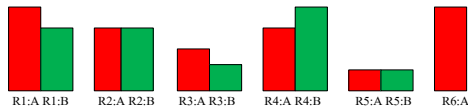
ภาพที่ 2: วิชาที่ใช้ในการจัดสอบ

จากภาพที่ 2 แสดงตัวอย่างข้อมูลจำนวนนักศึกษาสอบแต่ละวิชาโดย **A ปกติ** แทนจำนวนนักศึกษาภาคปกติที่สอบวิชา **A**, **A พิเศษ** แทนจำนวนนักศึกษาโครงการพิเศษที่สอบวิชา **A**



ภาพที่ 3: แสดงตัวอย่างจำนวนนักศึกษาที่ใช้ในการจัดตารางสอบ

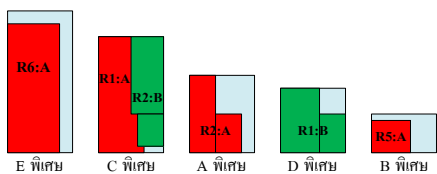
- จัดกลุ่มวิชาดังนี้ 1) โครงการพิเศษ-ภาคปกติ 2) เวลาสอบ ช่วงเช้า-ช่วงบ่าย 3) เรียงลำดับ นส. จากมากไปน้อย ดังภาพที่ 3
- เตรียมข้อมูลห้องสอบ



ภาพที่ 4: แสดงตัวอย่างห้องที่ใช้จัดสอบ

จากภาพที่ 4 แสดงตัวอย่างข้อมูลจำนวนที่นั่งสอบ โดย 1 ห้อง จัดเป็น part A และ part B สัญลักษณ์ แทนจำนวนที่นั่งสอบ ห้อง R5 มีจำนวนที่นั่ง part A และ part B

- จัดห้องสอบ พิจารณานักศึกษาโครงการพิเศษก่อนครั้งละ 1 วิชาและพิจารณาวิชาที่มีจำนวนนักศึกษามากก่อนตามลำดับ



ภาพที่ 5 : การจัดห้องลงวิชาสอบโดยพิจารณาห้องที่เล็กที่สุด

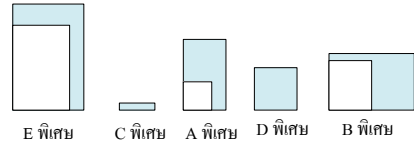
จากภาพที่ 5 เลือกจัดสอบ วิชา **E พิเศษ** โดยพิจารณาขั้นตอนวิธี A ดังนี้

A.1 พิจารณาห้องที่มีจำนวนนักศึกษาเท่ากับจำนวนที่นั่งสอบ กรณีนี้ไม่มีข้อมูลตรงตามเงื่อนไขพิจารณา **A.2**

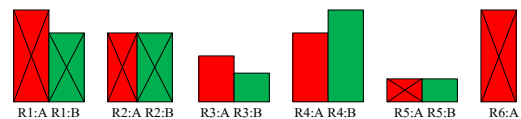
A.2 พิจารณาห้องที่เล็กที่สุดโดยมีจำนวนนักศึกษาน้อยกว่าจำนวนที่นั่งภายในห้อง กรณีนี้ไม่มีข้อมูลตรงเงื่อนไขพิจารณา **A.3**

A.3 เลือกห้องใหญ่ที่สุดและมีจำนวนนักศึกษามากกว่าจำนวนที่นั่งภายในห้อง

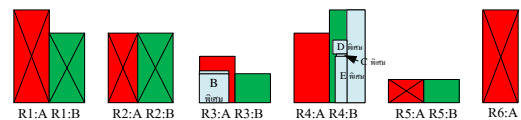
A.4 กลับไปทำข้อ A.1 ตามลำดับจนครบทุกวิชา จะมีวิชาที่มีจำนวนนักศึกษาที่ยังไม่จัดสอบ ดังภาพที่ 6 และสัญลักษณ์ แทนห้องที่ใช้จัดสอบแล้ว ดังภาพ ที่ 7



ภาพที่ 6 : จำนวนวิชาที่ยังไม่จัดสอบ



ภาพที่ 7 : ห้องสอบที่สามารถจัดสอบได้



ภาพที่ 8 : การจัดวิชาลงห้องสอบโดยพิจารณาห้องที่ใหญ่ที่สุด

จากภาพที่ 8 แสดงการจัดห้องสอบโดยพิจารณาวิชาที่ยังไม่จัดสอบแทนด้วย จากภาพที่ 7 เลือกห้องที่มีการใช้งานแล้ว, มีที่นั่งว่างและมีสถานะเดียวกับวิชาที่จะจัดสอบ (ภาคปกติ/โครงการพิเศษ) แทนด้วย **R3:A** **R3:B** จากภาพที่ 8 จึงพิจารณาห้องที่ยังไม่มีการจัดสอบ ซึ่งพิจารณาขั้นตอนวิธี **A.1, A.2, A.3, และ A.4** ดังภาพที่ 8

7. กรณีมีจำนวนนักศึกษายังไม่จัดสอบพิจารณาเงื่อนไขรองคือไม่สลับแถวในวิชาเดียวกัน ซึ่งพิจารณาขั้นตอนวิธี **A.1, A.2, A.3, และ A.4**

8. กรณีมีจำนวนนักศึกษายังไม่จัดสอบพิจารณาเงื่อนไขรองคือนักศึกษาภาคปกติและโครงการพิเศษสอบห้องเดียวกันแต่คิดภาระงานคุมสอบเป็นโครงการพิเศษ ซึ่งพิจารณาขั้นตอนวิธี **A.1, A.2, A.3, และ A.4** จะได้ผลการทดลองสรุปค่าใช้จ่ายทั้งหมดในการจัดตารางสอบ โดยอธิบายการทดลองการจัดสอบทั้งหมด 7 วิธี ดังนี้

1. การจัดสอบด้วยผู้เชี่ยวชาญ จัดสอบโดยคำนึงถึงห้องที่มีขนาดใหญ่ก่อน และพิจารณาครั้งละวิชา แต่ต้องใช้เวลาในการ

ตารางที่ 3 : แสดงค่าใช้จ่ายการจัดตารางสอบ 2 ภาคการศึกษา

ปีการศึกษา	จัดสอบด้วยผู้เชี่ยวชาญ	First-Fit Algorithm	Best Fit Algorithm	Best Fit Algorithm2	เลือก 2วิชาจำกัดห้อง (1-Bin-A)	เลือก 1วิชาไม่จำกัดห้อง (2-Bin-A)	เลือก 1วิชาจำกัดห้อง (3-Bin-A)
M 1/2555	44,616.76	69,907.18	71,044.38	73,481.10	46,316.25	47,982.18	41,054.10
F 1/2555	88,274.35	133,653.82	138,733.52	137,559.24	90,118.27	106,783.83	84,405.58
M 2/2555	37,293.76	53,462.15	55,521.62	56,779.30	35,644.98	43,800.59	33,608.10
F 2/2555	77,453.40	118,571.40	128,278.61	126,961.27	78,671.12	92,654.52	76,882.36

จัดสอบยาวนาน และไม่คำนึงถึงค่าใช้จ่ายในการจัดตารางสอบ

2. First-Fit Algorithm เป็นการจัดสอบที่นำห้องมาจัดสอบแต่ละวิชาที่จำนวนนักศึกษามากหรือเท่ากับจำนวนที่นั่ง และนำจำนวนนักศึกษาที่เหลือมาจัดสอบในห้องมีขนาดใหญ่กว่าหรือเท่ากับจำนวนนักศึกษาที่ต้องการ

3. Best Fit Algorithm เป็นการจัดสอบโดยการคำนึงถึงจำนวนที่นั่งสอบแต่ละห้องให้มีที่ว่างเหลือน้อยที่สุด โดยการพิจารณาห้องสอบทุกห้องทุกแถวก่อนเลือกจัดสอบ และเลือกจัดสอบครั้งละ 1 วิชา

4. Best Fit Algorithm2 (พิจารณาครั้งละ 2 วิชา) เป็นการจัดสอบโดยการคำนึงถึงจำนวนที่นั่งสอบแต่ละห้องให้มีที่ว่างเหลือน้อยที่สุด โดยการพิจารณาห้องสอบทุกห้องทุกแถวก่อนเลือกจัดสอบ และเลือกจัดสอบครั้งละ 2 วิชา

5. พิจารณาครั้งละ 2 วิชาและจำกัดห้อง (1-Bin-A) มีการคำนวณที่นั่งเพื่อเลือกอาคารที่เหมาะสมในการจัดสอบแล้วเหลือที่นั่งว่างน้อยที่สุด และแบ่งเงื่อนไขในการจัดสอบเป็น 2 เงื่อนไข คือเงื่อนไขหลักและเงื่อนไขรอง และจัดสอบครั้งละ 2 วิชา เพื่อให้เหลือที่นั่งว่างน้อยที่สุด

6. พิจารณาครั้งละ 1 วิชาและไม่จำกัดห้อง (2-Bin-A) แบ่งเงื่อนไขในการจัดสอบเป็น 2 เงื่อนไข คือเงื่อนไขหลักและเงื่อนไขรอง และจัดสอบครั้งละ 1 วิชา เพื่อให้เหลือที่นั่งว่างน้อยที่สุด

7. พิจารณาครั้งละ 1 วิชาจำกัดห้อง (3-Bin-A) มีการคำนวณที่นั่งเพื่อเลือกอาคารที่เหมาะสมในการจัดสอบแล้วเหลือที่นั่งว่างน้อยที่สุด และแบ่งเงื่อนไขในการจัดสอบเป็น 2 เงื่อนไข คือเงื่อนไขหลักและเงื่อนไขรอง และจัดสอบครั้งละ 1 วิชา เพื่อให้เหลือที่นั่งว่างน้อยที่สุด

3.3 การวัดประสิทธิภาพ

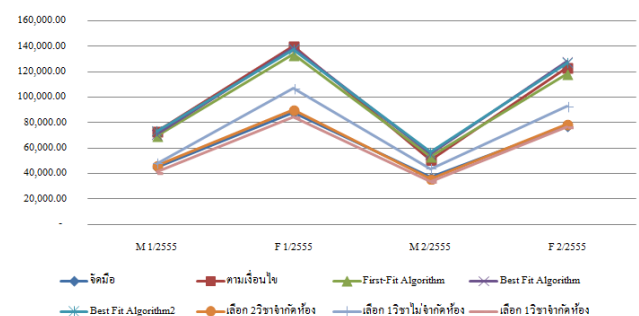
จากตารางที่ 3 แสดงผลการจัดตารางสอบด้วยวิธีการแบบจัดสอบด้วยผู้เชี่ยวชาญมีประสิทธิภาพสูงรองจากการจัดสอบ

ด้วยวิธี 3-Bin-A แต่อย่างไรก็ตามหากขาดบุคลากรที่ทำหน้าที่นี้เฉพาะจะทำให้ระบบงานล่าช้าหรือเกิดความผิดพลาดได้ ดังนั้นจึงได้คิดวิเคราะห์รูปแบบใหม่เพื่อนำมาใช้ในการจัดตารางสอบ นอกเหนือจากการพิจารณาที่นั่งว่างเหลือน้อยที่สุดแล้วยังพิจารณาจำนวนการใช้ห้องสอบ ซึ่งตัวแบบนี้พิจารณาถึงความเหมาะสมของอาคารเรียนที่ใช้จัดสอบแต่ภาคการศึกษาเพื่อใช้ทรัพยากรในการจัดสอบได้คุ้มค่าที่สุดเท่าที่จะทำได้ แสดงได้จากผลการจัดสอบด้วยวิธี 3-Bin-A ในคอลัมน์สุดท้าย

4. ผลการทดลอง

บทความนี้ได้ทำการทดลองจัดตารางสอบของคณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น ด้วยเทคนิคการบรรจุงานโดยใช้ข้อมูลการจัดตารางสอบจริง เป็นข้อมูลในการทดสอบกับตัวแบบที่ได้พัฒนาตามขั้นตอนวิธีการที่ได้ออกแบบเอาไว้

การคำนวณพลังงานไฟฟ้า ต้องนำพลังงานการใช้ไฟฟ้าแอร์และแสงสว่าง แล้วนำผลรวมไปคำนวณค่าใช้จ่าย โดยเฉลี่ยค่าพลังงานไฟฟ้า 1 หน่วย / 4 บาท แต่ถ้าใช้อาคาร SC.08 จะต้องคิดค่า Chiller หากเปิดใช้แอร์ 3 ชั่วโมง ใช้พลังงานไฟฟ้าประมาณ 299.53 KW แม้จะใช้ห้องมากหรือน้อย พลังงานไฟฟ้า Chiller มีค่าเท่าเดิม



ภาพที่ 9 : ผลการจัดตารางสอบ 2 ภาคการศึกษา

ในส่วนของการถูกต้องการจัดสอบด้วยเทคนิคการบรรจุงาน เป็นไปตามเงื่อนไขและมีประสิทธิภาพ แสดงผลการจัดตารางสอบได้ดังภาพที่ 9 ทั้งนี้ ผลการทดลองจะแสดงเพียง

บางส่วนเท่านั้น ผลการทดลองจะเห็นว่า วิธีการจัดตารางสอบด้วยวิธีด้วยเทคนิคการบรรจุงานสามารถแก้ปัญหาจัดตารางสอบได้อย่างมีประสิทธิภาพและเป็นไปตามเงื่อนไข

เมื่อพิจารณาในส่วนของค่าใช้จ่ายในการจัดสอบ จะเห็นว่าด้วยเทคนิคการบรรจุงานแบบ 3-Bin-A สามารถจัดตารางสอบโดยมีค่าใช้จ่ายพลังงานไฟฟ้าที่น้อยที่สุด แสดงดังภาพ 9 โดยแกน Y แทนค่าใช้จ่ายจากการคุมสอบแต่ภาคการศึกษาและแกน X คือ ภาคการศึกษาที่นำมาใช้เป็นข้อมูลทดสอบ

จากค่าใช้จ่ายจากการจัดสอบแต่ละภาคการศึกษา อาศัยหลักการพิจารณาโดยใช้ Bin Packing คือส่วนของการบรรจุและส่วนของพื้นที่ที่เหลือไป [12] โดยแบบ 3-Bin-A มีค่าใช้จ่ายในการจัดสอบน้อยที่สุด ทั้งค่าตอบแทนกรรมการคุมสอบ และค่าพลังงานไฟฟ้าที่ใช้ในการจัดห้องสอบ แสดงดังภาพที่ 9

5. สรุปผลและข้อเสนอแนะ

บทความนี้ได้นำอัลกอริทึมแบบใหม่สำหรับการจัดตารางสอบของคณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น ด้วยเทคนิคการบรรจุงานเพื่อแก้ปัญหาที่เกิดขึ้นจากวิธีการจัดตารางสอบแบบเดิมที่ใช้คนจัด ซึ่งในการทดลองได้ใช้ข้อมูลทดสอบเป็นข้อมูลการจัดตารางสอบของคณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น จากการทดลอง สรุปได้ว่า วิธีการจัดตารางสอบด้วยเทคนิคการบรรจุงานแบบ 3-Bin-A จะจัดวิชาสอบลงห้องสอบโดยเน้นให้มีค่าใช้จ่ายในการจัดสอบน้อยที่สุดต่างจากวิธีการจัดแบบใช้ผู้เชี่ยวชาญในการจัดที่เน้นจัดสอบโดยใช้ห้องสอบน้อยที่สุด ทำให้การจัดตารางสอบแบบแบบใช้ผู้เชี่ยวชาญ มีค่าใช้จ่ายในการจัดสอบ 2 ภาคการศึกษา มากกว่าแบบ 3-Bin-A 11,688.13 บาท

ในงานวิจัยนี้ไม่ได้พิจารณาไปถึงส่วนพื้นที่การจัดที่นั่งสอบ หากโปรแกรมสามารถวิเคราะห์พื้นที่ห้องสอบแล้วประมวลผลถึงจำนวนที่นั่งสอบที่สามารถจัดเพิ่มเติมได้ ก็จะเป็นวิธีการลดค่าใช้จ่ายในการจัดสอบ การนำข้อมูลส่วนนี้เข้ามารวมนในงานวิจัยทำให้ค่อนข้างลำบาก ในงานวิจัยในอนาคตจึงควรมีการเพิ่มศักยภาพในการทำงานของตัวแบบ โดยการเพิ่มข้อมูลในส่วนของการรายละเอียดพื้นที่ห้องสอบ เพื่อเป็นการเพิ่มประสิทธิภาพในการจัดตารางสอบโดยใช้พื้นที่ให้คุ้มค่าที่สุด

เอกสารอ้างอิง

- [1] วราญ อินทวงค์, “การจัดตารางสอบด้วยวิธีโอบีเจกต์เคมิ่ง” วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต, มหาวิทยาลัยรามคำแหง, 2553.
- [2] วณาลี แก้วพะนอ, สุทามาศ มนต์ริย์, “การจัดตารางสอบเพื่อให้คาบเวลาสอบเสร็จเร็วที่สุด” โครงการของนักศึกษาชั้นปีที่ 4, วิศวกรรมศาสตรบัณฑิต, มหาวิทยาลัยขอนแก่น, 2546.
- [3] อรุณี ปุกแก้ว, “การพัฒนาโปรแกรมจัดตารางสอนตารางสอบสำหรับคณะบริหารธุรกิจ สถาบันเทคโนโลยีราชมงคล วิทยาเขตพายัพ” วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต, มหาวิทยาลัยเชียงใหม่, 2546.
- [4] มนต์สวรรค์ ชันโท, “การพัฒนาระบบจัดตารางสอนตารางสอบกรณีศึกษา สถาบันราชภัฏกาญจนบุรี” สารนิพนธ์วิทยาศาสตรมหาบัณฑิต, มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2548.
- [5] นพดล บุรณ์กุล, “การจัดตารางสอบประจำภาคเรียนโดยใช้อัลกอริทึมแบบไฮบริด” สารนิพนธ์วิทยาศาสตรมหาบัณฑิต, มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2550.
- [6] ประกิจพันธ์ พัฒนศิริรัฐราช, “รูปแบบและการพัฒนาระบบสารสนเทศสำหรับการจัดตารางสอบ ระดับโรงเรียนมัธยมศึกษาในสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กระทรวงศึกษาธิการ” วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต, มหาวิทยาลัยราชภัฏเชียงราย, 2552.
- [7] สุรเดช ตรีวิทยากรานต์, “การจัดตารางสอบหัวข้อวิจัยโดยอาศัยเทคนิค Genetic Algorithms” วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต, มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี, 2540.
- [8] จิระเดช พลสวัสดิ์, ประภาส จงสถิตวัฒนา, “การแก้ปัญหาการจัดบรรจุผลิตภัณฑ์ในสามมิติโดยขั้นตอนวิธีพันธุกรรมดัดแปลง” จุฬาลงกรณ์มหาวิทยาลัย, 2546.
- [9] อติเรก แก้วมะหิงษ์, มารอง ผดุงสิทธิ์, “การจัดตารางงานในระบบเวลาจริงบนมัลติโพรเซสเซอร์โดยอาศัยแผนการอัดบรรจุแบบครอบคลุม” วารสารบัณฑิตวิทยาลัย บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, ฉบับที่ 2 มกราคม – ธันวาคม 2552.
- [10] สูดารัตน์ สุ่มมาตย์, “การใช้วิธีเชิงวิวิธคติเพื่อแก้ปัญหาการจัดเส้นทางยานพาหนะและการบรรจุ” วิทยานิพนธ์วิศวกรรมศาสตรมหาบัณฑิต มหาวิทยาลัยเกษตรศาสตร์, 2547.
- [11] ภาควิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยฟาร์อีสเทอร์น, “การจัดสรรเนื้อที่แบบพหุ” การจัดการหน่วยความจำหลัก (ออนไลน์). สืบค้นจาก http://bc.feu.ac.th/pichate/os_c/8chapter.8htm.
- [12] Majid Shakhshi-Niyaei, Fariborz Jolai, Jafar Razmi, “Extending Two Dimensional Bin Packing Problem: Consideration of Priority for Items”, Journal of Industrial and Systems Engineering Vol.3 No. 2 Summer 2009, pp. 72-83,2009.

การพัฒนาสื่อการเรียนรู้เรื่องการจราจรทางบกบนระบบปฏิบัติการแอนดรอยด์

The Development of Learning Media on the Topic of Land Traffic on Android Operating System

นัทธรียา สงัย (Nutthreeya Songyai)¹ และ วงศ ศรีอุไร (Wongkot Sriurai)²

^{1,2}ภาควิชาคณิตศาสตร์ สถิติ และคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยอุบลราชธานี

¹nutthreeya@gmail.com, ²scwongsr@ubu.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อ 1) พัฒนาสื่อการเรียนรู้ เรื่อง การจราจรทางบกบนระบบปฏิบัติการแอนดรอยด์ 2) หาประสิทธิภาพของสื่อการเรียนรู้ให้มีประสิทธิภาพตามเกณฑ์ 80/80 และ 3) ศึกษาความพึงพอใจของผู้ใช้สื่อการเรียนรู้ กลุ่มตัวอย่างที่ใช้ เป็นผู้ใช้ทั่วไปจำนวน 30 คน เครื่องมือที่ใช้ในการวิจัยได้แก่ 1) สื่อการเรียนรู้ เรื่อง การจราจรทางบกบนระบบปฏิบัติการแอนดรอยด์ 2) แบบทดสอบ และ 3) แบบประเมินความพึงพอใจ สถิติที่ใช้ ได้แก่ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน

ผลการวิจัยพบว่า สื่อการเรียนรู้ที่พัฒนาขึ้นมีประสิทธิภาพเท่ากับ 80.11/86.89 สูงกว่าเกณฑ์ที่กำหนดไว้ 80/80 และผู้ใช้มีความพึงพอใจต่อสื่อการเรียนรู้ ที่พัฒนาขึ้นในระดับมากที่สุด ($\bar{x} = 4.59$, $S.D. = 0.54$) ผลการวิจัยสรุปได้ว่า สื่อการเรียนรู้สามารถนำไปใช้งานได้

คำสำคัญ: สื่อการเรียนรู้ การจราจรทางบก ระบบปฏิบัติการแอนดรอยด์

Abstract

The purposes of this research were to 1) develop the learning media on the topic of land traffic on android operating system 2) study the effectiveness of the developed learning media to meet criteria of the efficiency 80/80 and 3) study the satisfaction of the users who learned by using the developed learning media. The samples were 30 users. The research instruments consisted of the learning media on the topic of land traffic on android operating system, an achievement test and a satisfaction questionnaire. The research statistics were used mean and standard deviation.

Results of the research were as follows: The efficiency of the learning media on the topic of land traffic on android operating system was 80.11/86.89, which was higher than the assigned criterion of 80/80. The users' satisfaction with learning through use of the learning media was the highest level. ($\bar{x} = 4.59$, $S.D. = 0.54$) In conclusion, the findings indicate that the learning media is usable and useful.

Keyword: Learning media , Land traffic , Android Operating System

1. บทนำ

ปัจจุบันสื่อการเรียนรู้มีหลากหลายรูปแบบ เช่น สื่อสิ่งพิมพ์ และสื่อเทคโนโลยี เป็นต้น โดยสื่อเหล่านี้มีความสำคัญต่อการเรียนรู้ของผู้เรียน ทำให้ผู้เรียนสามารถเรียนรู้ เข้าใจ เนื้อหาได้ง่ายขึ้น ซึ่งสื่อการเรียนรู้ที่ผลิตขึ้นด้วยโปรแกรมคอมพิวเตอร์มีความเหมาะสมที่จะนำมาใช้ในการนำเสนอ เนื่องจากสามารถสื่อสารได้ทั้งตัวอักษร ภาพนิ่ง ภาพเคลื่อนไหว วิดีโอ และเสียง ที่ผสมผสานกันได้อย่างลงตัว ทำให้สามารถเรียนรู้ได้อย่างมีประสิทธิภาพ [1] ในยุคนี้ความรู้เกี่ยวกับการจราจรทางบกถือได้ว่ามีความสำคัญต่อชีวิตประจำวันอย่างยิ่ง เนื่องจากมีผู้ใช้รถใช้ถนนเป็นจำนวนมาก คนส่วนใหญ่ใช้ยานพาหนะส่วนตัวในการเดินทางมากขึ้น ไม่ว่าจะเป็นการเดินทางโดยรถจักรยานยนต์หรือรถยนต์ [2] ซึ่งผู้ขับขี่ยานพาหนะบางคนยังไม่มีความรู้ความเข้าใจในเรื่องการจราจรทางบกมากเท่าที่ควร เพราะความรู้เรื่องการจราจรทางบกไม่ได้ถูกบัญญัติไว้ในหลักสูตรของการเรียนการสอนขั้นพื้นฐาน ทำให้ผู้ขับขี่ไม่สามารถที่จะนำความรู้ที่มีมาใช้ในการสอบขอใบอนุญาตขับขี่ได้ ซึ่งตามกฎหมายผู้ขับขี่ยานพาหนะทุกคนจะต้องมี

ใบอนุญาตฉบับที่ที่ออกโดยกรมการขนส่งจังหวัด ผู้ขับขี่ยานพาหนะจะต้องเข้ารับการอบรม ทำแบบทดสอบในภาคทฤษฎีและภาคปฏิบัติให้ถูกต้องตามเกณฑ์ จึงจะสามารถขอรับใบอนุญาตขับรถได้

ปัจจุบันสื่อที่ให้ความรู้ด้านการจราจรทางบกมีการนำเสนอในหลากหลายรูปแบบ เช่น ให้ความรู้ผ่านเว็บไซต์ แผ่นพับ และในรูปแบบหนังสือ เป็นต้น ซึ่งความรู้ที่ได้จากสิ่งเหล่านี้ เป็นเพียงการให้ข้อมูลในรูปแบบของตัวหนังสือและภาพประกอบ ยังไม่มีสื่อที่ช่วยให้ผู้เรียนสามารถเรียนรู้ได้ง่ายและมีความน่าสนใจ ทำให้ผู้ที่ขอสอบใบขับขี่บางคนที่มีความแตกต่างกัน เช่น มีอายุแตกต่างกัน และมีความสามารถในการรับรู้ที่ต่างกัน สิ่งเหล่านี้จึงเป็นอุปสรรคสำคัญที่ทำให้ผู้ขอรับใบอนุญาตขับขี่หลายคนสอบไม่ผ่าน

จากปัญหาดังกล่าวผู้วิจัยจึงได้พัฒนาสื่อการเรียนรู้การจราจรทางบกบนระบบปฏิบัติการแอนดรอยด์ โดยสื่อนี้สามารถใช้เป็นสื่อช่วยสำหรับการฝึกอบรมแก่ผู้ที่สนใจเข้าสอบขอใบอนุญาตขับขี่ให้มีความรู้ความเข้าใจเกี่ยวกับเนื้อหาที่ต้องใช้ในการทดสอบในภาคทฤษฎีและภาคปฏิบัติเพื่อขอรับใบอนุญาตขับขี่มากขึ้น ผู้เรียนสามารถที่จะเข้าศึกษาเนื้อหาของสื่อการเรียนรู้พื้นฐานการจราจรทางบกได้ทุกที่ทุกเวลา

2. วัตถุประสงค์ของงานวิจัย

2.1 เพื่อพัฒนาสื่อการเรียนรู้การจราจรทางบกบนระบบปฏิบัติการแอนดรอยด์

2.2 เพื่อหาประสิทธิภาพของสื่อการเรียนรู้การจราจรทางบกบนระบบปฏิบัติการแอนดรอยด์ ที่ผู้วิจัยพัฒนาขึ้น

2.3 เพื่อศึกษาความพึงพอใจของผู้ใช้สื่อการเรียนรู้การจราจรทางบกบนระบบปฏิบัติการแอนดรอยด์

3. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

3.1 ทฤษฎีที่เกี่ยวข้อง

3.1.1 กระบวนการพัฒนาสื่อการเรียนการสอนของ ADDIE Model

ADDIE Model [3] เป็นกระบวนการพัฒนาระบบการเรียนการสอน (Instructional System Design: ISD) ที่นิยมใช้กันโดยทั่วไปและสามารถนำมาใช้เป็นแนวทางในการพัฒนาสื่อ

การเรียนการสอนได้เป็นอย่างดี ADDIE เป็นรูปแบบการสอนที่ออกแบบขึ้นมา เพื่อใช้ในการออกแบบและพัฒนาระบบการเรียนการสอน ซึ่งเป็นที่ยอมรับกันโดยทั่วไปว่าสามารถนำไปใช้ออกแบบและพัฒนามาบทเรียนคอมพิวเตอร์ได้เป็นอย่างดี ซึ่ง ADDIE Model มีลำดับการพัฒนา 5 ขั้นตอน ประกอบด้วย การวิเคราะห์ (Analysis) การออกแบบ (Design) การพัฒนา (Development) การนำไปใช้ (Implement) และการประเมินผล (Evaluation)

3.1.2 การประเมินประสิทธิภาพสื่อการเรียนรู้

การหาประสิทธิภาพของสื่อการเรียนรู้ตามเกณฑ์มาตรฐาน E1/E2 [4] คือ การแสดงค่าประสิทธิภาพของกระบวนการระหว่างเรียน และค่าประสิทธิภาพรวมของผลสัมฤทธิ์ทางการเรียน ซึ่งนิยามประสิทธิภาพ E1/E2 มีดังนี้

E 1 หมายถึง ค่าร้อยละของคะแนนเฉลี่ยที่เกิดจากการทำกิจกรรมระหว่างเรียนจากชุดการสอนของผู้เรียน

E 2 หมายถึง ค่าร้อยละของคะแนนเฉลี่ยที่เกิดจากการทำแบบทดสอบหลังการเรียนของผู้เรียน

3.2 งานวิจัยที่เกี่ยวข้อง

วงษ์สุวัฒน์ โคงพิมาย [5] งานวิจัยนี้ได้พัฒนาและหาประสิทธิภาพบทเรียนคอมพิวเตอร์ช่วยสอนประเภท สถานการณ์จำลอง เรื่อง ความปลอดภัยในการขับขี่รถจักรยานยนต์ระดับชั้นมัธยมศึกษาปีที่ 4 เครื่องมือที่ใช้เก็บรวบรวมข้อมูลคือบทเรียนคอมพิวเตอร์ช่วยสอน กลุ่มตัวอย่างที่ทดลองใช้สื่อ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 4 โรงเรียนหนองกรดวัฒนา อำเภอด่านขุนทด จังหวัดนครราชสีมา ที่กำลังศึกษาในปีการศึกษา 2547 จำนวน 20 คน ผลการวิจัยพบว่าบทเรียนคอมพิวเตอร์ช่วยสอนที่พัฒนาขึ้นมีประสิทธิภาพ 80.62 / 81.66 เป็นไปตามเกณฑ์ที่กำหนดไว้ 80/80

สวียา สุรมณีและคณะ [6] งานวิจัยนี้มีวัตถุประสงค์เพื่อ 1) พัฒนามาบทเรียนคอมพิวเตอร์ช่วยสอนแบบจำลองสถานการณ์บนเครือข่ายอินเทอร์เน็ต เรื่องกฎหมายจราจรให้มีประสิทธิภาพตามเกณฑ์ 80/80 2) ศึกษาดัชนีประสิทธิผลของการเรียนรู้ด้วยบทเรียน 3) เปรียบเทียบผลสัมฤทธิ์ทางการเรียนก่อนเรียนและหลังเรียน และ 4) ศึกษาความพึงพอใจของนักเรียนที่มีต่อบทเรียน กลุ่มตัวอย่างที่ใช้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 4/10 โรงเรียนวาปีปทุม จังหวัดมหาสารคาม

จำนวน 45 คน เครื่องมือที่ใช้ในการวิจัยคือ บทเรียนคอมพิวเตอร์ช่วยสอน แบบประเมินคุณภาพบทเรียน แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน และแบบสอบถามความพึงพอใจ ผลการวิจัยพบว่า บทเรียนมีประสิทธิภาพเท่ากับ 84.44/84.15 สูงกว่าเกณฑ์ที่กำหนดไว้ 80/80 ค่าดัชนีประสิทธิผลของบทเรียนเท่ากับร้อยละ 73 นักเรียนมีผลสัมฤทธิ์ทางการเรียนหลังเรียนสูงกว่าก่อนเรียน และนักเรียนมีความพึงพอใจต่อบทเรียนที่พัฒนาขึ้นในระดับมากที่สุด

จากงานวิจัยข้างต้นพบว่า การพัฒนาสื่อการเรียนรู้ที่เกี่ยวกับการจราจรทางบกส่วนใหญ่เน้นนำเสนอในรูปแบบของข้อความ เสียงประกอบ รูปภาพ และมีภาพเคลื่อนไหวบางส่วน ทำให้ผู้เรียนไม่ได้ให้ความสนใจในสื่อการเรียนรู้เท่าที่ควร และเนื้อหาที่นำเสนอขึ้นยังไม่ครอบคลุมในการที่จะนำมาใช้ในการขอสอบใบอนุญาตขับขี่ สำหรับสื่อการเรียนรู้การจราจรบนระบบปฏิบัติการแอนดรอยด์ที่ผู้วิจัยพัฒนาขึ้นนี้ จะนำเสนอข้อมูลทั้งข้อความ ภาพนิ่ง ภาพเคลื่อนไหว วิดีโอ พร้อมเสียงอธิบายประกอบ ทำให้สื่อการเรียนรู้มีความน่าสนใจ เนื้อหาที่นำเสนอมีความครอบคลุมในส่วนของกรนำไปใช้ในการทดสอบขอรับใบอนุญาตขับขี่ โดยผู้เรียนสามารถศึกษาบทเรียนได้ทุกที่ทุกเวลา

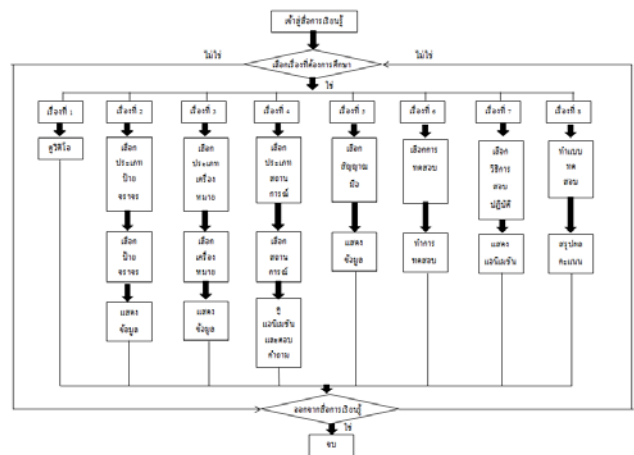
4. วิธีดำเนินการวิจัย

4.1 ประชากรและกลุ่มตัวอย่าง

ประชากรคือ ผู้ใช้ทั่วไปที่ทดสอบเพื่อขอรับใบอนุญาตขับขี่ ในจังหวัดอุบลราชธานี จำนวน 60 คน โดยได้คัดเลือกกลุ่มตัวอย่างมาจำนวน 30 คน ซึ่งได้มาจากวิธีการสุ่มโดยบังเอิญ

4.2 เนื้อหาที่ใช้ในการวิจัย

ผู้วิจัยได้แบ่งเนื้อหาออกเป็น 8 เรื่อง ได้แก่ 1) การขอรับใบอนุญาตขับรถต่างๆ 2) ป้ายจราจร 3) เครื่องหมายที่เจ้าพนักงานแสดงออกด้วยมือ 4) กฎจราจรที่ควรทราบ 5) สัญญาณที่เจ้าพนักงานแสดงออกด้วยมือ 6) การทดสอบสมรรถภาพทางร่างกาย 7) วิธีการสอบปฏิบัติ 8) ตัวอย่างแนวข้อสอบภาคทฤษฎีเพื่อขอรับใบอนุญาตขับรถ ดังแสดงในภาพที่ 1



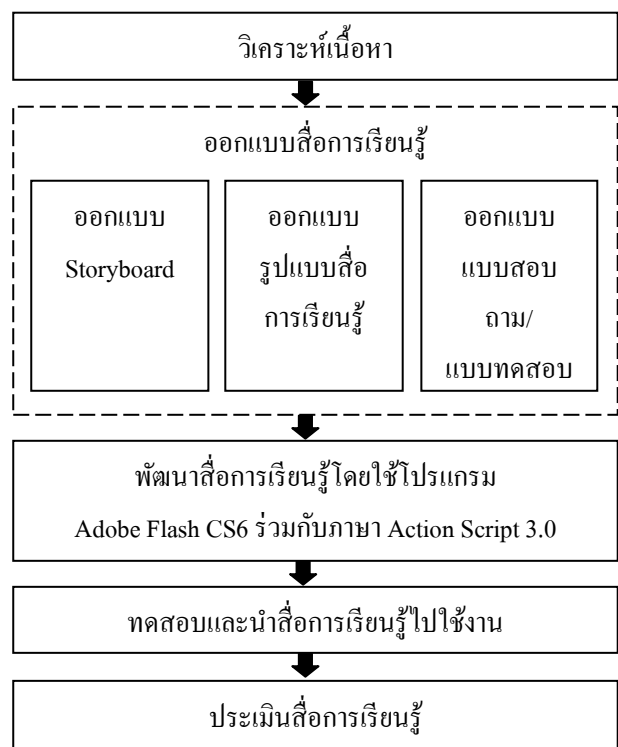
ภาพที่ 1: แผนผังสื่อการเรียนรู้การจราจรทางบก

4.3 เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยประกอบด้วย 1) สื่อการเรียนรู้ เรื่อง การจราจรทางบกบนระบบปฏิบัติการแอนดรอยด์ 2) แบบทดสอบ และ 3) แบบประเมินความพึงพอใจของผู้ใช้สื่อการเรียนรู้ที่ผู้วิจัยพัฒนาขึ้น

4.4 ขั้นตอนการดำเนินการวิจัย

ผู้วิจัยนำแนวคิดการสร้างสื่อการเรียนรู้ โดยใช้แบบจำลอง ADDLE Model มาใช้ในการพัฒนาสื่อการเรียนรู้ ดังแสดงในภาพที่ 2



ภาพที่ 2: ขั้นตอนการดำเนินการวิจัย

จากภาพที่ 2 สามารถอธิบายรายละเอียดในแต่ละขั้นตอนได้ดังนี้

1) วิเคราะห์เนื้อหา

ขั้นตอนนี้เป็นการวิเคราะห์เนื้อหาก่อนที่จะทำการออกแบบบทเรียน ผู้วิจัยได้กำหนดวัตถุประสงค์ คือ ให้ผู้ที่สนใจสอบใบอนุญาตขับขี่รถยนต์และรถจักรยานยนต์ มีความรู้ความเข้าใจเกี่ยวกับพื้นฐานการจราจรทางบก สำหรับเนื้อหาที่ใช้ในการสอบใบอนุญาตขับขี่รถยนต์และรถจักรยานยนต์ จะมีเนื้อหาหลักๆ ได้แก่ ความรู้ทั่วไปของการขอรับใบอนุญาตขับขี่รถต่างๆ ป้ายจราจรและเครื่องหมายจราจรบนพื้นทางในแต่ละประเภท กฎจราจรต่าง ๆ สัญญาณที่เข้าพนักงานแสดงออกด้วยมือ ซึ่งเนื้อหาเหล่านี้จะใช้ในการทดสอบในภาคทฤษฎี ส่วนการทดสอบในภาคปฏิบัตินั้นผู้สอบใบอนุญาตขับขี่รถจะต้องขับรถในสนามสอบและปฏิบัติตามให้ถูกต้องตามเกณฑ์ที่ตั้งไว้ก่อน จึงจะสามารถรับใบอนุญาตขับขี่รถได้

2) ออกแบบสื่อการเรียนรู้

เมื่อทำการวิเคราะห์เนื้อหาที่เกี่ยวข้องแล้ว ขั้นตอนนี้เป็นกรออกแบบ Storyboard และออกแบบรูปแบบของการนำเสนอสื่อการเรียนรู้ รวมทั้งออกแบบแบบสอบถามและแบบทดสอบก่อนและหลังศึกษาสื่อการเรียนรู้ จำนวน 30 ข้อ เพื่อให้ผู้ใช้ได้ทำแบบทดสอบ

3) พัฒนาสื่อการเรียนรู้

ผู้วิจัยพัฒนาสื่อการเรียนรู้ โดยใช้โปรแกรม Adobe Flash CS6 ร่วมกับภาษา Action Script 3.0

4) ทดสอบและนำสื่อการเรียนรู้ไปใช้งาน

ขั้นตอนนี้เป็นการทดสอบสื่อการเรียนรู้โดยใช้วิธีการทดสอบแบบ Black Box Testing และได้นำสื่อการเรียนรู้ที่พัฒนาขึ้นไปให้ผู้ใช้งานจำนวน 30 คนได้ทดสอบการใช้งาน

5) ประเมินสื่อการเรียนรู้

หลังจากที่ทำการทดสอบและได้ทดลองนำสื่อการเรียนรู้ไปใช้งานแล้ว ได้ดำเนินการประเมินสื่อโดยใช้แบบประเมินความพึงพอใจจากผู้ที่ได้ทดลองนำสื่อไปใช้งาน

5. ผลการวิจัย

5.1 ผลที่ได้จากการพัฒนาสื่อการเรียนรู้

ผู้วิจัยได้พัฒนาสื่อการเรียนรู้การจราจรทางบกบนระบบปฏิบัติการแอนดรอยด์ ซึ่งประกอบด้วยเนื้อหาทั้งหมด 8 เรื่อง โดยแสดงเป็นภาพตามลำดับดังนี้



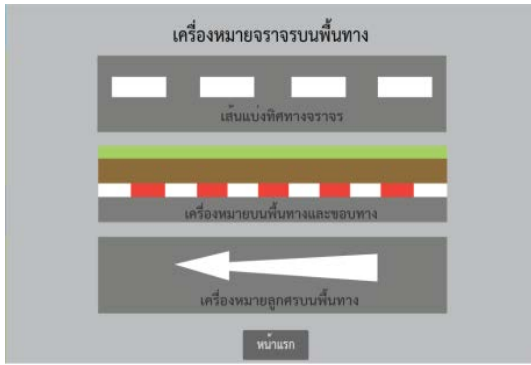
ภาพที่ 3: หน้าแรกของสื่อการเรียนรู้



ภาพที่ 4: เรื่องที่ 1 การขอรับใบอนุญาตขับรถต่างๆ



ภาพที่ 5: เรื่องที่ 2 ป้ายจราจร



ภาพที่ 6: เรื่องที่ 3 เครื่องหมายจราจรบนพื้นทาง



ภาพที่ 10: เรื่องที่ 7 วิธีการสอบปฏิบัติ



ภาพที่ 7: เรื่องที่ 4 กฎจราจรที่ควรทราบ



ภาพที่ 11: เรื่องที่ 8 ตัวอย่างข้อสอบภาคทฤษฎีเพื่อขอรับใบอนุญาตขับรถ



ภาพที่ 8: เรื่องที่ 5 สัญญาณที่เจ้าพนักงานแสดงออกด้วยมือ



ภาพที่ 9: เรื่องที่ 6 การทดสอบสมรรถภาพทางร่างกาย

5.2 ผลการประเมินประสิทธิภาพสื่อการเรียนรู้

ผู้วิจัยได้นำสื่อการเรียนรู้การจราจรทางบกบนระบบปฏิบัติการแอนดรอยด์ ไปทดลองใช้กับผู้ใช้งานไปจำนวน 30 คน เพื่อหาประสิทธิภาพของสื่อการเรียนรู้ ผลปรากฏดังตารางที่ 1

ตารางที่ 1 : ประสิทธิภาพของสื่อการเรียนรู้การจราจรทางบกบนระบบปฏิบัติการแอนดรอยด์ ตามเกณฑ์ 80/80 (E1/E2)

เกณฑ์	ค่าประสิทธิภาพ
E1	80.11
E2	86.89

จากตารางที่ 1 สรุปได้ว่า สื่อการเรียนรู้การจราจรทางบกบนระบบปฏิบัติการแอนดรอยด์ที่พัฒนาขึ้น มีประสิทธิภาพตามเกณฑ์ (80.11/86.89)

5.3 ผลการประเมินความพึงพอใจ

กำหนดระดับความพึงพอใจของการใช้สื่อการเรียนรู้ดังนี้ [7]

- 4.51 – 5.00 หมายถึง มากที่สุด
- 3.51 – 4.50 หมายถึง มาก
- 2.51 – 3.50 หมายถึง ปานกลาง
- 1.51 – 2.50 หมายถึง น้อย
- 1.00 – 1.50 หมายถึง น้อยที่สุด

ผู้วิจัยได้ศึกษาความพึงพอใจของผู้ใช้สื่อการเรียนรู้อยู่โดยใช้แบบสอบถามความพึงพอใจจากผู้ใช้งาน 30 คน ผลปรากฏดังตารางที่ 2

ตารางที่ 2: ผลประเมินความพึงพอใจของผู้ใช้สื่อการเรียนรู้อาจารย์ทางบกบนระบบปฏิบัติการแอนดรอยด์

รายการประเมิน	\bar{X}	S.D.	การแปลผล
1. ด้านเนื้อหา	4.63	0.53	มากที่สุด
2. ด้านภาพ ภาษา และเสียง	4.55	0.53	มากที่สุด
3. ด้านตัวอักษรและสี	4.58	0.57	มากที่สุด
4. ด้านการออกแบบแบบทดสอบ	4.58	0.53	มากที่สุด
5. ด้านการจัดการสื่อการเรียนรู้อ	4.63	0.58	มากที่สุด
เฉลี่ยรวม	4.59	0.54	มากที่สุด

จากตารางที่ 2 สรุปได้ว่าความพึงพอใจของผู้ใช้สื่อการเรียนรู้อ มีความพึงพอใจอยู่ในระดับมากที่สุด

6. สรุปผลการวิจัย

ผลของการวิจัย พบว่า สื่อการเรียนรู้อาจารย์ทางบกบนระบบปฏิบัติการแอนดรอยด์ ที่พัฒนาขึ้นมีประสิทธิภาพตามเกณฑ์ 80/80 เมื่อนำสื่อการเรียนรู้อไปทดลองใช้กับผู้ใช้ทั่วไป จำนวน 30 คน พบว่าผู้ใช้สื่อการเรียนรู้อส่วนใหญ่มีความพึงพอใจอยู่ในระดับมากที่สุด การวิจัยครั้งนี้ สรุปได้ว่า สื่อการเรียนรู้อพื้นฐานการจราจรทางบกบนระบบปฏิบัติการแอนดรอยด์ที่พัฒนาขึ้น มีประสิทธิภาพและสามารถนำไปใช้ประกอบในการให้ความรู้หรือใช้ประกอบการเรียนการสอนและการอบรมได้

7. กิตติกรรมประกาศ

ขอขอบพระคุณ ท่านอดุลย์ โครงการแข่งขันพัฒนาโปรแกรมคอมพิวเตอร์แห่งประเทศไทย ครั้งที่ 16 จากศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ

เอกสารอ้างอิง

- [1] วันสนันท์ ยศช้อย. สื่อมัลติมีเดียเพื่อการเรียนการสอน. (2555). เข้าถึงได้จาก : <http://wanussanun.wordpress.com/category/บทความ> [วันที่ค้นข้อมูล : 5 กุมภาพันธ์ 2557]
- [2] สุรศักดิ์ ลัทธสิ. กฎหมายเกี่ยวกับจราจรทางบก. (2553). เข้าถึงได้จาก : http://bpsmakom.org/BP_School/Social/Law1/Law-Street-traffic.htm [วันที่ค้นข้อมูล : 5 กุมภาพันธ์ 2557]
- [3] วิภา จันทร์จันเมฆ. การออกแบบระบบการสอน. (2555). เข้าถึงได้จาก : <http://com544144009unit5.blogspot.com/> [วันที่ค้นข้อมูล : 5 กุมภาพันธ์ 2557]
- [4] มนตรี แยมกลีกร. เกณฑ์ประสิทธิภาพในงานวิจัยและพัฒนาสื่อการสอน. วารสารศึกษาศาสตร์ ปีที่ 19 ฉบับที่ 1 เดือนตุลาคม 2550-มกราคม 2551. หน้า 1-16.
- [5] วงษ์สุวัฒน์ ไค่งพิมาย. การพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนประเภทสถานการณ์จำลอง เรื่อง ความปลอดภัยในการขับขี่รถจักรยานยนต์ ระดับชั้นมัธยมศึกษาปีที่ 4. [วิทยานิพนธ์ปริญญาครุศาสตรมหาบัณฑิต]. มหาวิทยาลัยราชภัฏมหาสารคาม.2547.
- [6] สวียา สุรมณี พิศุทธา อารีราษฎร์ วิลัน จุมปาแผ่และวิทยา อารีราษฎร์ .การพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนแบบจำลองสถานการณ์บนเครือข่ายอินเทอร์เน็ต เรื่อง กฎหมายจราจร.วารสารมหาวิทยาลัยราชภัฏมหาสารคาม ปีที่ 4 ฉบับที่ 1 : มกราคม - เมษายน 2553 : 153-160.
- [7] ราตรี นันทสุนนธ์. 2554. การวิจัยในชั้นเรียนและการวิจัยพัฒนาการเรียนการสอน. หน้า 133 – 136.

การพัฒนาสื่อการเรียนรู้เรื่องเซลล์และโครโมโซมด้วยเทคโนโลยีโลกเสมือนผสมผสานโลกจริง

Development of Learning Media in Topics of Cell and Chromosome using Augmented Reality Technology

ณัฐ คิชเจริญ (Nadh Ditcharoen)¹ จูติกพร ประครองญาติ (Thitikorn Prakrongyat)¹

นลพรรณ ประลอบพันธุ์ (Nonlapan Praloppan)¹ และสุภาพร พรไตร (Supaporn Porntrai)²

¹ภาควิชาคณิตศาสตร์ สถิติและคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยอุบลราชธานี

²ภาควิชาวิทยาศาสตร์ชีวภาพ คณะวิทยาศาสตร์ มหาวิทยาลัยอุบลราชธานี

scnadhdi@ubu.ac.th, hottie_cpb@hotmail.com, moobile32@gmail.com, sportrai@gmail.com

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาสื่อการเรียนรู้รายวิชาชีววิทยา เรื่อง เซลล์และโครโมโซม โดยใช้เทคโนโลยีโลกเสมือนผสมผสานโลกจริง ที่สามารถซ้อนภาพสามมิติให้ไปอยู่บนภาพที่เห็นจริงผ่านกล้อง ซึ่งใช้งานบนระบบปฏิบัติการแอนดรอยด์ ผู้ใช้งานสามารถมองเห็น โมเดลที่จำลององค์ประกอบลักษณะการเคลื่อนไหวและการทำงานของเซลล์และโครโมโซม ในรูปแบบแอนิเมชันสามมิติ จำนวน 21 โมเดล เครื่องมือที่ใช้ในการพัฒนาประกอบด้วย โปรแกรม Autodesk Maya โปรแกรม Unity และโปรแกรม Photoshop การตรวจสอบความถูกต้องของโมเดลทำโดยการวิเคราะห์ความสอดคล้องของเนื้อหา (ค่า IOC) และศึกษาความพึงพอใจจากผู้เชี่ยวชาญของมหาวิทยาลัยอุบลราชธานี จำนวน 4 คน ผลการพัฒนาและทดสอบเบื้องต้นพบว่า ความสอดคล้องของเนื้อหา มีค่าเฉลี่ยเท่ากับ 0.70 และค่าเฉลี่ยของความพึงพอใจ มีค่าเท่ากับ 4.32 (เต็ม 5) สื่อการเรียนรู้นี้ช่วยกระตุ้นความสนใจและเพิ่มความเข้าใจให้แก่ผู้เรียนตอบสนองการเรียนรู้ของผู้เรียนได้ทุกที่ทุกเวลา

คำสำคัญ: สื่อการเรียนรู้ ระบบปฏิบัติการแอนดรอยด์ เซลล์โครโมโซม

Abstract

The objective of this research was to develop a media for learning biology in topics of cell and chromosome using Augmented Reality, in which 3D virtual models are integrated into a 3D real environment through a camera in real time. The media was designed for running on an android device displaying 21 models of compartments

and movement of cell and chromosome in form of 3D animations. It was developed using Unity, Autodesk Maya and Adobe Photoshop. All 3D animations were examined by 4 experts of Ubon Ratchathani University in terms of item-objective congruence (IOC) and users' satisfaction. The preliminary development and test results showed that the index of IOC was 0.70 and the average of users' satisfaction was 4.32 (from 5). The learning media is able to motivate and enhance learners' interest and understanding. It is also met the demand of learners in learning anywhere anytime.

Keyword: Learning Media, Android OS, Cell, Chromosome

1. บทนำ

การเรียนการสอนในรายวิชาชีววิทยา เรื่องเซลล์และโครโมโซมจำเป็นต้องอาศัยจินตนาการช่วยทำความเข้าใจ ทั้งนี้เพราะเซลล์และโครโมโซมไม่สามารถมองเห็นด้วยตาเปล่า เนื่องจากมีขนาดเล็กมาก ซึ่งปัจจุบันการเรียนรู้เรื่องนี้จะศึกษาผ่านหนังสือที่มีรูปภาพประกอบแบบสองมิติ ซึ่งหากจะดูรายละเอียดถึงเรื่องสีหรือขนาดก็ต้องซื้อตำราในราคาแพงขึ้น อย่างไรก็ตามก็ยังมีสื่อการเรียนรู้ด้านนี้อยู่บ้าง ทั้งที่จัดทำในรูปแบบวีดิโอ หรือแอนิเมชันในรูปแบบของไฟล์แฟลช (Flash Animation) แต่ก็ยังไม่แพร่หลายนัก บริษัทลานเกียร์ (www.lamgeartech.com) ร่วมกับสถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (สสวท.) [1] ได้พัฒนาสื่อการเรียนรู้วิทยาศาสตร์หลายเรื่องโดยการประยุกต์ใช้เทคโนโลยี

โลกเสมือนผสมผสานโลกจริง (Augmented Reality: AR) ซึ่งเป็นการซ้อนภาพสามมิติที่อยู่ในโลกเสมือน (Virtual World) ให้ไปอยู่บนภาพที่เห็นจริงผ่านกล้องดิจิทัล หรือเว็บแคม ที่ผู้เรียนสามารถมองเห็นได้แบบ 360 องศา อย่างไรก็ตามยังไม่มีสื่อการเรียนรู้อะไรที่ในเรื่องเซลล์และโครโมโซม ดังนั้นผู้วิจัยจึงได้พัฒนาสื่อการเรียนรู้อะไรที่ในเรื่องเซลล์และโครโมโซมโดยใช้เทคโนโลยี AR ที่สามารถใช้งานได้บนระบบปฏิบัติการแอนดรอยด์ เพื่อให้อาจารย์หรือครูผู้สอนรายวิชาชีววิทยาได้ใช้ประกอบการเรียนการสอน และให้ผู้สนใจเรียนรู้อะไรที่ในเรื่องเซลล์และโครโมโซมใช้ศึกษาได้ทุกที่ทุกเวลา

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 เทคโนโลยีโลกเสมือนผสมผสานโลกจริง

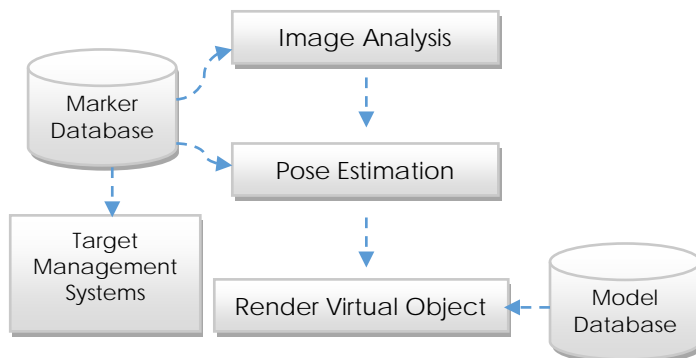
เทคโนโลยีโลกเสมือนผสมผสานโลกจริง (Augmented Reality) หรือ AR เป็นเทคโนโลยีที่นำเอาภาพกราฟิกของคอมพิวเตอร์ทั้งในรูปแบบที่เป็นสามมิติหรือวิดีโอ มาซ้อนทับเข้ากับฉากหลังซึ่งเป็นภาพสองมิติ แสดงผลแบบเรียลไทม์ (Real Time) AR นั้นมีการใช้งานอย่างแพร่หลายในประเทศสหรัฐอเมริกา และประเทศญี่ปุ่น โดยเริ่มต้นแนวคิดนี้ตั้งแต่ปี 2533 และเป็นรูปธรรมในปี 2540 เป็นแนวคิดการผสมผสานสิ่งที่คอมพิวเตอร์แสดงผลด้วยตัวละครเสมือนกับพื้นหลังซึ่งเป็นโลกแห่งความจริง บนพื้นฐานของหลักการแกนสามมิติ กระบวนการภายในของเทคโนโลยีนี้ ประกอบด้วย 4 ส่วน (ดังแสดงในภาพที่ 1) [2] คือ

1) การวิเคราะห์ภาพ (Image Analysis) เป็นขั้นตอนของการค้นหามาร์คเกอร์จากภาพที่ได้มาจากกล้องแล้วทำการสืบค้นจากฐานข้อมูล (Marker Database) ที่มีการเก็บข้อมูลขนาดและรูปแบบของมาร์คเกอร์ที่ถูกออกแบบไว้

2) การคำนวณตำแหน่งเชิง 3 มิติ (Pose Estimation) ของมาร์คเกอร์เทียบกับกล้อง แสดงในรูปแบบเมตริกซ์ที่ระบุความสัมพันธ์ระหว่างตำแหน่งของกล้องและตำแหน่งของมาร์คเกอร์

3) กระบวนการสร้างภาพโมเดลสามมิติ (Render Virtual Object) เทคโนโลยี AR เป็นการเพิ่มข้อมูล (โมเดลสามมิติ) เข้าไปในภาพที่ได้จากกล้อง ณ ตำแหน่งที่ตรวจพบจากขั้นตอนที่ 1) โดยใช้ค่าตำแหน่งจากขั้นตอนที่ 2)

4) ระบบจัดการเป้าหมาย (Target Management) เป็นระบบที่ใช้ในการจัดการมาร์คเกอร์ ที่มาพร้อมกับการทำงานของ Vuforia (www.vuforia.com) ซึ่งเป็นเทคโนโลยีการแสดงผลภาพจำลองเสมือนจริงของบริษัท Qualcomm (รายละเอียดในหัวข้อ 3.3)



ภาพที่ 1 : การทำงานของเทคโนโลยี AR

ตัวอย่างการประยุกต์ใช้เทคโนโลยีโลกเสมือนผสมผสานโลกจริงในการพัฒนาสื่อการเรียนรู้อะไรที่ในวิชาศาสตร์ของบริษัทลานเกียร์ เทคโนโลยี จำกัด ร่วมกับสถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (สสวท.) พัฒนาสื่อการเรียนรู้อะไรที่ในชุด “AR Learning Gears – Science Series” ซึ่งประกอบด้วย 6 เรื่อง ได้แก่ โครงสร้างโลก ระบบสุริยะ บันทึกลงโลก แผ่นดินไหว โครงสร้างอะตอม และการจมน้ำและการลอย ตัวอย่างหนังสือดังแสดงในภาพที่ 2



(ก) เรื่อง โครงสร้างโลก (ข) ระบบสุริยะจักรวาล

ภาพที่ 2 : ตัวอย่างสื่อการเรียนรู้อะไรที่ในวิชาศาสตร์ของบริษัทลานเกียร์ ที่มา:

www.larngeartech.com

2.2 งานวิจัยที่เกี่ยวข้อง

วรรณวิภา อนุกุลสวัสดิ์ และคณะ [3] พบว่ารายวิชาชีววิทยาต้องอาศัยการทำความเข้าใจในเนื้อหามากกว่าวิชาอื่น เพราะเป็นวิชาที่มีเนื้อหาห่างไกลจากการดำรงชีวิตจริงของผู้เรียน มีข้อจำกัดหลายอย่างที่เรายังไม่สามารถทำการศึกษาหรือหาการอ้างอิงที่ถูกต้องเหมาะสมมาใช้ในการเรียนการสอน

ได้ อีกทั้งยังเป็นเนื้อหาวิชาที่สามารถมองภาพหรือสร้างจินตนาการได้ยากต่อการเข้าใจ มักเกิดความเข้าใจผิดและเกิดความคลาดเคลื่อนเกี่ยวกับข้อถูกต้องในเรื่องต่างๆ ของลักษณะเฉพาะหรือรูปร่างของส่วนต่างๆ ในสิ่งมีชีวิตและไม่มีชีวิต ที่เป็นส่วนหนึ่งในบทเรียนวิชาชีววิทยา

พรชัย จันทรสุกแสง [4] กล่าวถึงหนังสือสามมิติและเทคโนโลยีโลกเสมือนผสมผสานโลกจริง ว่ายังไม่เป็นที่แพร่หลายในปัจจุบัน ไทยยังต้องนำเข้าจากต่างประเทศ และมีราคาแพง มีการนำมาใช้บ้างในงานทางด้านโฆษณาเท่านั้น แต่ในแง่ของการศึกษายังไม่มีการนำมาใช้แต่อย่างใด เพราะการผลิตกราฟิกสามมิติต้องลงทุนสูงมากจึงยังไม่มีเอกชนไทยรายใดกล้าลงทุน แต่มั่นใจว่าในอนาคตจะได้เห็นสื่อการเรียนรู้ประเภทนี้ออกมาให้เห็นมากขึ้น เพราะหนังสือสามมิติจะสามารถกระตุ้นให้เด็กเกิดการเรียนรู้ได้มากยิ่งขึ้น และยังเป็นเครื่องมือช่วยกระตุ้นให้เด็กเห็นว่าการเรียนเป็นเรื่องที่สนุก ไม่ใช่เรื่องน่าเบื่ออีกต่อไป

รัชพล ชนานวงษ์ [5] พัฒนาคือการเรียนรู้เรื่องการจมและการลอยด้วยเทคโนโลยี AR เป็นหนึ่งในชุดสื่อการเรียนรู้ “AR Learning Gears – Science Series” ที่ สสวท. ร่วมกับบริษัทลานเกียร์พัฒนาขึ้น ทั้งนี้ยังสนับสนุนว่าเทคโนโลยีโลกเสมือนผสมผสานโลกจริงจะทำให้สามารถมองเห็นถึงลักษณะการเคลื่อนไหวและการทำงานภายในและภายนอกของเซลล์และโครโมโซมได้อย่างเสมือนจริง

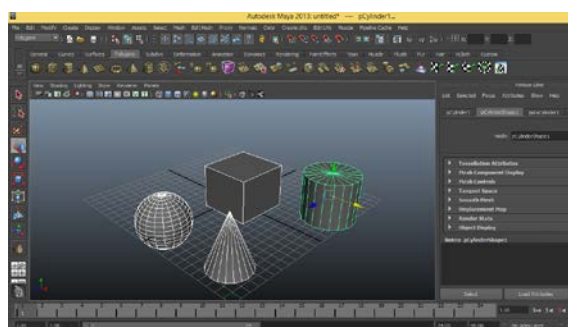
3. การออกแบบและพัฒนาระบบ

ในการพัฒนาคือการเรียนรู้ ผู้วิจัยแบ่งการทำงานออกเป็น 3 ส่วน คือ 1) การออกแบบและพัฒนาโมเดล 2) การออกแบบและพัฒนามาร์คเกอร์ และ 3) การพัฒนาแอปพลิเคชันรายละเอียดดังต่อไปนี้

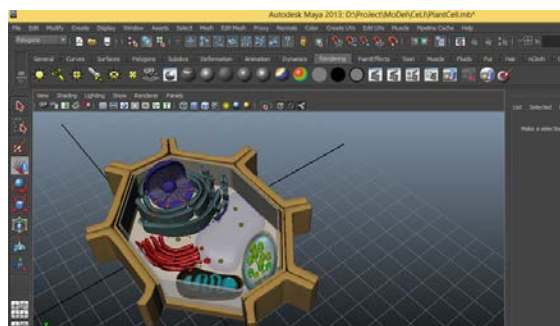
3.1 การออกแบบและพัฒนาโมเดล

ผู้วิจัยใช้โปรแกรม Autodesk Maya (Student Version) ในการออกแบบและพัฒนาโมเดลสามมิติ ซึ่ง Maya เป็นโปรแกรมที่ได้รับความนิยมอย่างแพร่หลาย เพราะมีความยืดหยุ่นในการใช้งานสูง สามารถปรับรูปแบบของโมเดล ได้อย่างอิสระ อีกทั้งยังสามารถนำเอาโมเดลที่ผ่านการพัฒนาแล้วมาพัฒนาต่อยอดได้ (Reusability) การปั้น โมเดลสาม มิติของ เซลล์และ

โครโมโซม ผู้วิจัยใช้การขึ้นรูปด้วยเครื่องมือสำเร็จรูปที่มีอยู่ในโปรแกรม Maya การขึ้นรูปจะมีทั้งเริ่มจากรูปทรงสี่เหลี่ยม ทรงกลม ทรงกรวยและทรงกระบอก ดังภาพที่ 3 โดยสามารถปรับแต่งวัตถุ (Object) เพื่อให้ได้รูปแบบตามที่กำหนด ดังภาพที่ 4 ซึ่งเป็นภาพโมเดลของเซลล์พืช จากนั้นทำการเคลื่อนไหวโมเดลในส่วนที่ได้กำหนดไว้ แล้วจึงทำการ Export ไฟล์ออกจากโปรแกรมโดยใช้สกุลไฟล์ .fbx เพื่อนำไปใช้ในขั้นตอนพัฒนาแอปพลิเคชันต่อไป



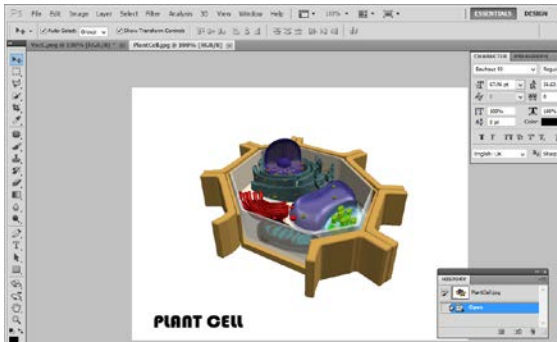
ภาพที่ 3 : การเริ่มออกแบบโมเดลด้วยรูปทรงต่างๆ



ภาพที่ 4 : การออกแบบโมเดลด้วยโปรแกรม Maya

3.2 การออกแบบและพัฒนามาร์คเกอร์

ผู้วิจัยใช้รูปภาพที่ได้จากการเรนเดอร์ภาพสามมิติด้วยโปรแกรม Maya มาพัฒนาเป็นมาร์คเกอร์ โดยการตกแต่งเพิ่มเติมด้วยโปรแกรม Photoshop ทำให้ได้มาร์คเกอร์ที่มีความหมายสัมพันธ์กับรูปแบบของโมเดล ซึ่งทำให้ผู้ใช้เข้าใจได้ง่ายขึ้น ดังภาพที่ 5 แสดงการออกแบบมาร์คเกอร์สำหรับโมเดลเซลล์พืช (Plant Cell)

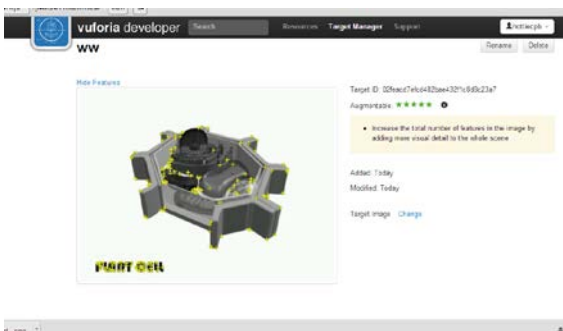


ภาพที่ 5 : การออกแบบและพัฒนามาร์คเกอร์

3.3 การออกแบบและพัฒนาแอปพลิเคชัน

การพัฒนาแอปพลิเคชันที่ใช้งานบนระบบปฏิบัติการแอนดรอยด์ โดยใช้โปรแกรม Unity 3D ร่วมกับ Vuforia (www.vuforia.com) เพื่อเชื่อมโยงมาร์คเกอร์ที่พัฒนาในข้อ 3.2 กับโมเดลในข้อ 3.1 ให้แท็บเล็ตหรือสมาร์ทโฟนระบบปฏิบัติการแอนดรอยด์สามารถอ่านค่ามาร์คเกอร์ผ่านกล้องแล้วแสดงผลโมเดลสามมิติที่ตรงกันได้ กระบวนการเริ่มจากการอัปโหลดภาพมาร์คเกอร์ที่ได้ในข้อ 3.2 ไปยังฐานข้อมูลของ Vuforia ดังภาพที่ 6 จากนั้นทำการดาวน์โหลดมาร์คเกอร์ดังกล่าวมาใช้ในโปรแกรม Unity ดังภาพที่ 7

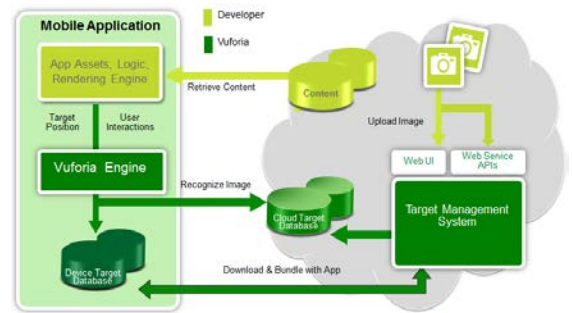
การทำงานร่วมกับโปรแกรม Unity ทำโดยใช้ Vuforia Engine (ภาพที่ 8) ดึงเอาข้อมูลของมาร์คเกอร์ใน Unity ซึ่งใช้คำว่า Target ที่ดึงจาก Device Target Database หรือ Cloud Target Database การจัดการในส่วนนี้กระทำผ่าน Target Management System ที่สามารถใช้งานได้บนเว็บไซต์ www.vuforia.com



ภาพที่ 6 : มาร์คเกอร์ที่ทำการอัปโหลดเรียบร้อยแล้ว



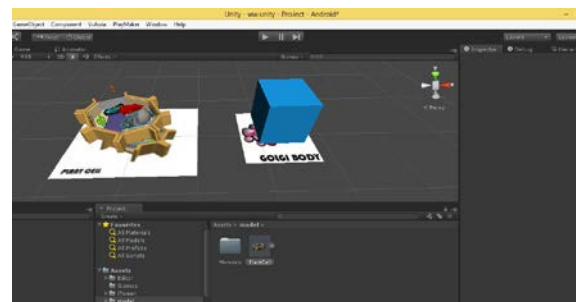
ภาพที่ 7 : การดาวน์โหลดมาร์คเกอร์มาใช้งาน



ภาพที่ 8 : กระบวนการทำงานของ Vuforia

ที่มา : <http://gravityjack.com/>

เมื่อทำการติดตั้งมาร์คเกอร์เสร็จแล้วจะใช้โปรแกรม Unity เชื่อมโยงมาร์คเกอร์กับโมเดลสามมิติ ดังภาพที่ 9 จากนั้นจึง Export ออกมาเป็นไฟล์นามสกุล .apk เพื่อให้สามารถนำไปติดตั้งบนระบบปฏิบัติการแอนดรอยด์ได้



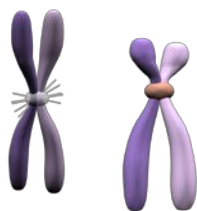
ภาพที่ 9 : การพัฒนาโปรแกรมบน Unity

4. ผลการพัฒนาและทดสอบระบบ

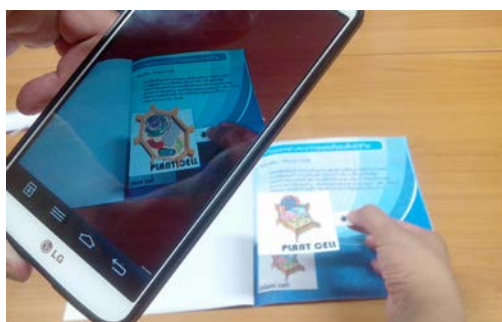
สื่อการเรียนรู้ เรื่องเซลล์และโครโมโซม ประกอบด้วยโมเดลสามมิติจำนวน 21 โมเดล ตัวอย่างโมเดลเกี่ยวกับโครงสร้างเซลล์แสดงดังภาพที่ 10 และโมเดลโครโมโซมแสดงดังภาพที่ 11 โมเดลทั้งหมดสามารถมองเห็นได้แบบ 360 องศา ผลการทดสอบโดยการใช้งานผ่านสมาร์ทโฟน บนระบบปฏิบัติการแอนดรอยด์ แสดงดังภาพที่ 12



ภาพที่ 10 : ตัวอย่างโมเดลเรื่องเซลล์



ภาพที่ 11 : ตัวอย่างโมเดลเรื่องโครโมโซม



ภาพที่ 12 : การทดสอบโดยใช้งานผ่านสมาร์ทโฟน

การประเมินประสิทธิภาพของสื่อการเรียนรู้ แบ่งออกเป็น 2 ส่วน ได้แก่ 1) การวิเคราะห์ความสอดคล้องของเนื้อหา (ค่า IOC) และ 2) การศึกษาความพึงพอใจของผู้ใช้งาน

ผู้วิจัยได้วิเคราะห์ความสอดคล้องของเนื้อหา (ค่า IOC : Index of Item-Objective Congruence) [6] จากสูตร ดังนี้

$$IOC = \frac{\sum R}{N}$$

โดยที่ $\sum R$ คือ ผลรวมของคะแนนผู้เชี่ยวชาญแต่ละคน และ N คือจำนวนผู้เชี่ยวชาญทั้งหมด

ผู้เชี่ยวชาญประกอบด้วย อาจารย์ผู้สอนประจำภาควิชาวิทยาศาสตร์ชีวภาพ คณะวิทยาศาสตร์ มหาวิทยาลัยอุบลราชธานี จำนวน 4 คน ผลการคำนวณค่า IOC มีค่าเฉลี่ยเท่ากับ 0.70 ซึ่งบ่งชี้ว่า โมเดลที่พัฒนาขึ้นสอดคล้องกับเนื้อหา และสามารถสื่อความหมายได้อย่างถูกต้อง ทั้งนี้ผู้เชี่ยวชาญมีความเห็นตรงกันในประเด็นที่ “สื่อการเรียนรู้นี้ทำให้ผู้เรียน

ตื่นเต้น เร้าใจ ทำให้เห็นรูปร่างที่แท้จริงของโครงสร้างต่างๆ ทำให้มีมุมมองและความเข้าใจเกี่ยวกับเซลล์และโครโมโซมกว้างขึ้น”

นอกจากนี้ยังได้ศึกษาความพึงพอใจต่อการใช้งานจากผู้เชี่ยวชาญดังกล่าวด้วยแบบสอบถาม แบ่งออกเป็น 3 ด้าน ได้แก่ 1) คุณภาพด้านเนื้อหา 2) คุณภาพด้านสื่อ และ 3) ความพึงพอใจในการใช้งาน เกณฑ์คะแนนแบ่งเป็น 5 ระดับ ตามมาตราส่วนประเมินค่าของลิเคิร์ต (Likert Scale) [6] ผลการวิเคราะห์ค่าเฉลี่ยของแบบสอบถาม แสดงดังตารางที่ 1

ตารางที่ 1 : ผลการวิเคราะห์แบบสอบถาม

คุณภาพ	ร้อยละของระดับคะแนนการให้ข้อมูล					ระดับเฉลี่ย
	มากที่สุด	มาก	ปานกลาง	น้อย	น้อยที่สุด	
ค่าคะแนนด้านเนื้อหา	(5)	(4)	(3)	(2)	(1)	4.29
ด้านสื่อ	35.71	57.14	7.14	0	0	4.18
การใช้งาน	25.00	67.86	7.14	0	0	4.18
	55.00	45.00	0	0	0	4.55
	ระดับเฉลี่ย					4.32
	เฉลี่ยร้อยละ					86.32

จากตารางที่ 1 ค่าเฉลี่ยของความพึงพอใจโดยรวมของทั้ง 3 ด้าน มีค่าเท่ากับ 4.32 คิดเป็น 86.32% ผู้เชี่ยวชาญมีความเห็นเพิ่มเติมว่า “สื่อที่พัฒนาขึ้นจะทำให้การเรียนรู้ชีววิทยาน่าสนใจมากขึ้น”

5. สรุปและอภิปรายผล

การพัฒนาสื่อการเรียนรู้เรื่องเซลล์และโครโมโซม โดยการนำเทคโนโลยีโลกเสมือนผสานโลกจริง (AR) มาประยุกต์ใช้ในการเรียนการสอนรายวิชาชีววิทยานี้ ถูกออกแบบให้ใช้งานบนระบบปฏิบัติการแอนดรอยด์ทั้งสมาร์ทโฟนและแท็บเล็ตซึ่งผู้ใช้งานสามารถหาซื้อได้สะดวกในราคาประหยัด มาร์คเกอร์ออกแบบให้อ่านค่าได้ง่ายและรวดเร็วด้วยกล้องและสื่อความหมายให้ผู้ใช้งานเข้าใจได้ง่ายด้วยภาพที่สร้างจากโมเดลสามมิติ การพัฒนาแอปพลิเคชันบนมือถือโดยใช้ Vuforia ช่วยให้สามารถอ่านค่ามาร์คเกอร์ได้อย่างแม่นยำในระยะเฉลี่ยไม่เกิน 80 เซนติเมตร นอกจากนี้ผู้วิจัยยังพัฒนาเว็บไซต์รวบรวม

ไฟล์ติดตั้งและมาร์คเกอร์เพื่ออำนวยความสะดวกให้ผู้เรียนได้ดาวน์โหลดแอปพลิเคชันไปใช้งานได้ทุกที่ตลอดเวลา

จากการทดสอบและประเมินผล พบว่าสื่อการเรียนรู้ช่วยสร้างความสนใจและดึงดูดใจผู้เรียน ผู้สอนสามารถนำสื่อนี้ไปใช้ขยายความรู้ (Elaborate) ให้ผู้เรียนได้พัฒนาความเข้าใจในเนื้อหาได้กว้างและลึกยิ่งขึ้น สามารถช่วยแก้ปัญหาความเข้าใจคลาดเคลื่อน (Misconception) ได้ดี ในอนาคตอันใกล้เทคโนโลยีโลกเสมือนผสมผสานโลกจริงนี้อาจเข้าไปมีส่วนร่วมในกระบวนการเรียนรู้แบบร่วมมือ (Collaborative Learning) หรือการเรียนรู้แบบอื่นๆ ที่สอดคล้องกับการเรียนรู้ในศตวรรษที่ 21

6. กิตติกรรมประกาศ

ขอกราบขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร.อรัญญา พิมพ์มงคล ผู้ช่วยศาสตราจารย์ ดร.ถาวร สุภาพรม และ ดร.สุทธนา ปลอดสมบุญ อาจารย์ประจำภาควิชาวิทยาศาสตร์ชีวภาพ คณะวิทยาศาสตร์ มหาวิทยาลัยอุบลราชธานี ที่กรุณาเป็นผู้เชี่ยวชาญประเมินประสิทธิภาพของสื่อการเรียนรู้ และแนะนำแนวทางสำหรับการพัฒนาสื่อการเรียนรู้ให้ตรงกับความต้องการและสามารถนำไปใช้ในการเรียนการสอนได้อย่างถูกต้องสมบูรณ์

เอกสารอ้างอิง

- [1] thumbsup, “ข่าวประชาสัมพันธ์: บริษัทลานเกียร์เทคโนโลยีเปิดตัวผลิตภัณฑ์สื่อการเรียนรู้ด้วยเทคโนโลยี AR” [Online] เข้าถึงได้จาก: <http://thumbsup.in.th/2012/08/lamgeartech-ar/> (วันที่สืบค้นข้อมูล: 10 กุมภาพันธ์ 2557)
- [2] พนิดา ตันศิริ, “โลกเสมือนผสมผสานโลกจริง (Augmented reality),” *วารสารนักบริหาร มหาวิทยาลัยกรุงเทพ* ปีที่ 30 ฉบับที่ 2 เม.ย.-มิ.ย. 2553 หน้า 169-175.
- [3] วรณวิภา อนุกุลสวัสดิ์, จิระวรรณ เกษสิงห์ และธีราพร อนันตะเศรษฐกุล, “แนวคิดเรื่องเซลล์และโครงสร้างเซลล์ของนักเรียนชั้นมัธยมศึกษาปีที่ 5,” *การประชุมทางวิชาการของมหาวิทยาลัยเกษตรศาสตร์ ครั้งที่ 50*, กรุงเทพฯ, 2555, หน้า 96-104.
- [4] พรชัย จันทรสุกแสง, “เรื่องหนังสือ 3 มิติ จากไบลานสู่จิตติคอล” [Online] เข้าถึงได้จาก: http://www.truelookpanya.com/true/blog_diary_detail.php?diary_id=3970. (วันที่สืบค้นข้อมูล: 10 กุมภาพันธ์ 2557)
- [5] รัชพล ธนานางค์, “สื่อเสริมการเรียนรู้ โลกเสมือนผสมผสานโลกจริง (Augmented Reality) เรื่องการจมน้ำจืด,” สถาบันส่งเสริมการ

สอนวิทยาศาสตร์และเทคโนโลยี สาขาวิทยาศาสตร์มัธยมศึกษา, [Online] เข้าถึงได้จาก:

http://secondsci.ipst.ac.th/index.php?option=com_content&view=article&id=336:armedia&catid=19:2009-05-04-05-01-56&Itemid=34 (วันที่สืบค้นข้อมูล: 10 กุมภาพันธ์ 2557).

- [6] มนต์ชัย เทียนทอง, *สถิติและวิธีการวิจัยทางเทคโนโลยีสารสนเทศ*. กรุงเทพฯ : สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2548.

การพัฒนาสื่อมัลติมีเดียแบบมีปฏิสัมพันธ์เรื่องการทดลองวิทยาศาสตร์บนระบบปฏิบัติการแอนดรอยด์ สำหรับนักเรียนชั้นประถมศึกษาปีที่ 2

The Development of Interactive Multimedia on the Topic of Science Experiment on Android Operating System for Grade 2 Students

อติพร ปานพุ่ม (Atiporn Panpum)¹ และวงกต ศรีอุไร (Wongkot Sriurai)²

^{1,2}ภาควิชาคณิตศาสตร์ สถิติ และคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยอุบลราชธานี

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อ 1) พัฒนาสื่อมัลติมีเดียแบบมีปฏิสัมพันธ์ เรื่องการทดลองวิทยาศาสตร์บนระบบปฏิบัติการแอนดรอยด์ สำหรับนักเรียนชั้นประถมศึกษาปีที่ 2 2) หาประสิทธิภาพของสื่อมัลติมีเดียแบบมีปฏิสัมพันธ์ให้มีประสิทธิภาพตามเกณฑ์ 80/80 และ 3) ศึกษาความพึงพอใจของผู้ใช้สื่อมัลติมีเดียแบบมีปฏิสัมพันธ์ กลุ่มตัวอย่างที่ใช้เป็นนักเรียนชั้นประถมศึกษาปีที่ 2 โรงเรียนบ้านศรีไค จังหวัดอุบลราชธานี จำนวน 30 คน เครื่องมือที่ใช้ในการวิจัยได้แก่ สื่อมัลติมีเดียแบบมีปฏิสัมพันธ์เรื่องการทดลองวิทยาศาสตร์บนระบบปฏิบัติการแอนดรอยด์ สำหรับนักเรียนชั้นประถมศึกษาปีที่ 2 แบบทดสอบ และ แบบประเมินความพึงพอใจ สถิติที่ใช้ได้แก่ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน ผลการวิจัยพบว่าสื่อมัลติมีเดียแบบมีปฏิสัมพันธ์ที่พัฒนาขึ้นมีประสิทธิภาพเท่ากับ 81.00/89.75 สูงกว่าเกณฑ์ที่กำหนดไว้ 80/80 และผู้ใช้มีความพึงพอใจต่อสื่อมัลติมีเดียแบบมีปฏิสัมพันธ์ที่พัฒนาขึ้น ในระดับมาก ($\bar{X} = 4.50, S.D. = 0.75$)

คำสำคัญ: มัลติมีเดียแบบมีปฏิสัมพันธ์ การทดลองวิทยาศาสตร์ ระบบปฏิบัติการแอนดรอยด์

Abstract

The purposes of this research were to 1) develop the interactive multimedia on the topic of science experimentation on android operating system for Grade 2 students 2) study the effectiveness of the developed interactive multimedia to meet criteria of the efficiency 80/80 and 3) study the satisfaction of the students who learned by using the developed interactive multimedia. The samples were 30 Grade 2 students from Ban Sri

Khai School in Ubon Ratchathani province. The research instruments consisted of the interactive multimedia on the topic of science experimentation on android operating system for Grade 2 students, an achievement test and a satisfaction questionnaire. The research statistics were used mean and standard deviation. Results of the research were as follows: The efficiency of the interactive multimedia was 81.00/89.75., which was higher than the assigned criterion of 80/80. The students' satisfaction with learning through use of the interactive multimedia was a high level. ($\bar{X} = 4.50, S.D. = 0.75$)

Keywords: Interactive Multimedia, Science Experiment, Android Operating System

1. บทนำ

ปัจจุบันการเรียนการสอนวิชาวิทยาศาสตร์ในระดับประถมศึกษา เน้นการทำกิจกรรมการเรียนการสอน โดยวิธีการปฏิบัติจริงตามหลักสูตรปีพ.ศ. 2551.ศ. [1] เนื้อหาในบทเรียนจึงเป็นเนื้อหาที่เน้นการทดลองเป็นหลักเพื่อให้ผู้เรียนได้เรียนรู้จากการปฏิบัติจริง นอกจากจะเรียนรู้จากบทเรียนในตำราเท่านั้น จากเนื้อหาที่เน้นการทดลองและการปฏิบัติของวิชาวิทยาศาสตร์จึงทำให้โรงเรียนจะต้องใช้งบประมาณในการจัดซื้ออุปกรณ์ทางการทดลองทางวิทยาศาสตร์ที่จำเป็น แต่ในปัจจุบันพบว่ายังมีบางโรงเรียนที่เกิดปัญหาการขาดแคลนอุปกรณ์การเรียนรู้อุปกรณ์ในวิชาวิทยาศาสตร์ และครูผู้สอนมีจำนวนน้อย ไม่เพียงพอต่อการจัดการเรียนการสอนแบบการปฏิบัติ เพราะการเรียนการสอนแบบนี้จำเป็นอย่างมากที่จะต้องใช้ให้ครูผู้สอนคอยควบคุมดูแลอย่างใกล้ชิดเนื่องจากวัสดุอุปกรณ์บางชิ้นอาจทำอันตรายกับผู้เรียนได้ จึงทำให้การเรียนการสอน

วิชาวิทยาศาสตร์ในโรงเรียนประถมศึกษาชั้นยังไม่มีความพร้อมพอที่จะรองรับการเรียนรู้ที่เพิ่มประสิทธิภาพ

จากปัญหาการขาดแคลนวัสดุอุปกรณ์การทดลองในวิชาวิทยาศาสตร์ จึงทำให้การเรียนการสอนในห้องเรียนนั้นเป็นการเรียนการสอนเชิงทฤษฎี ซึ่งการเรียนการสอนในส่วนนี้ทำการเรียนการสอนได้ง่ายกว่า ปัจจุบันมีการนำสื่อการเรียนการสอนต่างๆมาใช้สอนในวิชาวิทยาศาสตร์ เช่น โปสเตอร์ แผนภาพ และการ์ตูนแอนิเมชัน เพื่อให้ผู้เรียนเกิดความสนใจและเรียนเข้าใจได้ง่ายขึ้น สำหรับการเรียนการสอนระดับชั้นประถมศึกษาชั้น ปัจจุบันมีโครงการแท็บเล็ตพีซีเพื่อการศึกษาไทยจึงมีการผลิตสื่อประเภทการ์ตูนแอนิเมชัน ให้ผู้เรียนได้เรียนรู้ด้วยตนเองบนแท็บเล็ต ทำให้ผู้เรียนเกิดความสนใจในบทเรียนมากยิ่งขึ้น

เนื่องจากสื่อการ์ตูนแอนิเมชันสามารถดึงดูดความสนใจของผู้เรียนในวัยเยาว์ได้ จึงทำให้ปัจจุบันมีการพัฒนาการ์ตูนแอนิเมชันเพื่อใช้ในการเรียนการสอนมีจำนวนมากขึ้น และเริ่มมีการนำมาใช้อย่างแพร่หลายทั้งบนคอมพิวเตอร์และบนอุปกรณ์เคลื่อนที่ ซึ่งการพัฒนาเพื่อให้ใช้งานบนอุปกรณ์เคลื่อนที่ได้ นั้นกำลังเป็นที่นิยมอย่างมาก โดยเฉพาะระบบปฏิบัติการแอนดรอยด์ที่มีการใช้งานอย่างแพร่หลายในเวลาที่ผ่านมา เพราะมีราคาที่ถูก มีแอปพลิเคชันที่หลากหลาย และมีผู้พัฒนาแอปพลิเคชันจำนวนมาก

จากปัญหาที่กล่าวมาข้างต้นนั้น ปัญหาในเรื่องของการขาดแคลนวัสดุอุปกรณ์ทางการเรียนรู้วิชาวิทยาศาสตร์จำเป็นต้องมีการแก้ปัญหา ผู้วิจัยจึงได้คิดวิธีการแก้ปัญหาโดยการใช้วิธีการผลิตสื่อที่สามารถให้ผู้เรียนได้ทำการทดลองและสังเกตผลการทดลองได้บนอุปกรณ์เคลื่อนที่ที่ผู้เรียนมีอยู่แล้วนั่นคือแท็บเล็ต โดยการพัฒนาสื่อมัลติมีเดียการทดลองวิชาวิทยาศาสตร์ให้อยู่ในรูปแบบที่เป็นการ์ตูนแอนิเมชันที่สามารถโต้ตอบกับผู้เรียนได้ ซึ่งจะช่วยเพิ่มความน่าสนใจให้แก่ผู้เรียนและผู้เรียนได้มีส่วนร่วมในการทดลอง

2. วัตถุประสงค์ของงานวิจัย

2.1 เพื่อพัฒนาสื่อมัลติมีเดียแบบมีปฏิสัมพันธ์เรื่องการทดลองวิทยาศาสตร์บนระบบปฏิบัติการแอนดรอยด์

สำหรับนักเรียนชั้นประถมศึกษาปีที่ 2

2.2 เพื่อหาประสิทธิภาพของสื่อมัลติมีเดียแบบมีปฏิสัมพันธ์ที่ผู้วิจัยพัฒนาขึ้น

2.3 เพื่อศึกษาความพึงพอใจของผู้ใช้สื่อมัลติมีเดียแบบมีปฏิสัมพันธ์

3. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

3.1 ทฤษฎีที่เกี่ยวข้อง

3.1.1 สื่อมัลติมีเดีย [2] คือ การนำองค์ประกอบของสื่อชนิดต่างๆ มาผสมผสานเข้าด้วยกันเช่นตัวอักษร ภาพนิ่ง ภาพเคลื่อนไหว เสียง และ วิดีโอ โดยผ่านกระบวนการทางระบบคอมพิวเตอร์ เพื่อสื่อความหมายกับผู้ใช้และถ้าผู้ใช้สามารถควบคุมสื่อให้นำเสนอออกมาตามต้องการได้ เรียกว่า สื่อมัลติมีเดียปฏิสัมพันธ์ การปฏิสัมพันธ์ของผู้ใช้สามารถทำได้โดยผ่านทางคีย์บอร์ด หรือเมาส์ เป็นต้น

3.1.2 การประเมินประสิทธิภาพสื่อมัลติมีเดีย

การหาประสิทธิภาพของสื่อตามเกณฑ์มาตรฐาน E1/E2 [3] คือ การแสดงค่าประสิทธิภาพของกระบวนการระหว่างเรียน และค่าประสิทธิภาพรวบยอดของผลสัมฤทธิ์ทางการเรียน ซึ่งนิยามประสิทธิภาพ E1/E2 มีดังนี้

E1 หมายถึง ค่าร้อยละของคะแนนเฉลี่ยที่เกิดจากการทำกิจกรรมระหว่างเรียนจากชุดการสอนของผู้เรียน

E2 หมายถึง ค่าร้อยละของคะแนนเฉลี่ยที่เกิดจากการทำแบบทดสอบหลังการเรียนของผู้เรียน

3.2 งานวิจัยที่เกี่ยวข้อง

เนรมิต สุดชนะ [4] ได้พัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนกลุ่มสาระการเรียนรู้วิทยาศาสตร์ เรื่องระบบสุริยะ สำหรับชั้นประถมศึกษาปีที่ 4 งานวิจัยนี้ได้ศึกษาผลของการใช้บทเรียนคอมพิวเตอร์ช่วยสอน โดยมีการนำสื่อคอมพิวเตอร์ช่วยสอนมาทดลองใช้กับกลุ่มผู้เรียน ผลการวิจัยพบว่า ผู้เรียนมีความก้าวหน้าทางการเรียนเพิ่มขึ้น ร้อยละ และผู้เรียนมี 67 ความพึงพอใจต่อบทเรียนคอมพิวเตอร์ช่วยสอนอยู่ในระดับมาก

วรพรรณ บุคคีด้วง [5] ได้ทำการศึกษาผลของการใช้บทเรียนคอมพิวเตอร์มัลติมีเดีย กลุ่มสาระวิทยาศาสตร์ เรื่องสารในชีวิตประจำวัน ชั้นมัธยมศึกษาปีที่ 1 โดย งานวิจัยนี้มีการ

พัฒนาบทเรียนคอมพิวเตอร์มัลติมีเดีย และเปรียบเทียบผลสัมฤทธิ์ทางการเรียนของผู้เรียนก่อนและหลังการนำบทเรียนคอมพิวเตอร์เข้ามาใช้ ผลการวิจัยพบว่า ผู้เรียนมีผลสัมฤทธิ์ทางการเรียนที่สูงขึ้นเมื่อนำบทเรียนคอมพิวเตอร์มัลติมีเดียเข้ามาใช้งาน

จากงานวิจัยข้างต้นสรุปได้ว่า การนำสื่อการเรียนการสอนเข้ามาใช้ในการเรียนการสอนช่วยให้ผู้เรียนมีความก้าวหน้าทางการเรียนเพิ่มขึ้น

สำหรับงานวิจัยที่ได้จัดทำขึ้นนี้เป็นงานวิจัยที่มีความแตกต่างจากการ์ตูนแอนิเมชันทั่วไปคือ ผู้เรียนสามารถโต้ตอบกับตัวการ์ตูนผ่านทางกรรสัมพันธ์หน้าจอแท็บเล็ต ซึ่งจะทำให้ผู้เรียนเกิดความสนใจมากขึ้นกว่าการนั่งดูการ์ตูนเพียงอย่างเดียวและยังเน้นให้ผู้เรียนได้ทดลองด้วยตนเอง จะต่างจากสื่อทั่วไปที่จะเน้นให้ความรู้แต่จะไม่เน้นปฏิบัติ

4. วิธีดำเนินการ

4.1 ประชากรและกลุ่มตัวอย่าง

ประชากรคือ นักเรียนชั้นประถมศึกษาปีที่ 2 จาก 3 โรงเรียน ได้แก่ โรงเรียนบ้านแมคคำลือชา โรงเรียนบ้านแฮหนามแท่ง และโรงเรียนบ้านศรีไค รวมทั้งหมดจำนวน 70 คน กลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 2 โรงเรียนบ้านศรีไค จำนวน 30 คน การเลือกกลุ่มตัวอย่างใช้วิธีเลือกแบบเจาะจง

4.2 เนื้อหาที่ใช้ในการวิจัย

ผู้วิจัยได้คัดเลือกเนื้อหาการทดลองวิทยาศาสตร์ทั้งหมดจำนวน 4 หน่วยการเรียนรู้ ซึ่งประกอบด้วย 8 การทดลองได้แก่

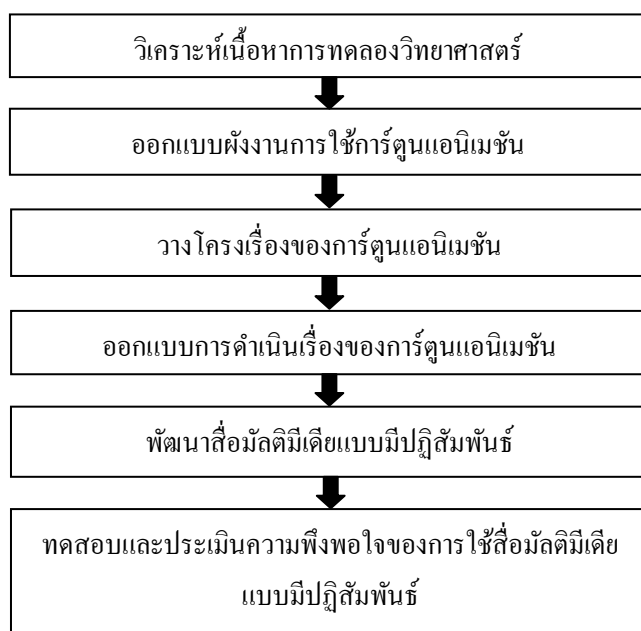
1. การทดลองการดูดซึมน้ำของวัสดุต่างๆ
2. การทดลองการดูดแม่เหล็กของวัตถุต่างๆ
3. การทดลองการดูดและการผลักของแม่เหล็ก 2 ตัว
4. การทดลองการเกิดไฟฟ้าสถิตในวัตถุต่างๆ
5. การทดลองการเกิดแรงไฟฟ้าระหว่างวัตถุ 2 ชนิด
6. การทดลองการต่อวงจรไฟฟ้า
7. การทดลองฉนวนและตัวนำไฟฟ้า
8. การทดลองการอุ้มน้ำของดิน 3 ชนิด (ดินร่วน ดินทราย และดินเหนียว)

4.3 เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยประกอบด้วย 1) สื่อมัลติมีเดียแบบมีปฏิสัมพันธ์เรื่อง การทดลองวิทยาศาสตร์ บนระบบปฏิบัติการแอนดรอยด์ สำหรับนักเรียนชั้นประถมศึกษาปีที่ 2 2) แบบทดสอบ และ 3) แบบประเมินความพึงพอใจของผู้ใช้สื่อมัลติมีเดียแบบมีปฏิสัมพันธ์ที่ผู้วิจัยพัฒนาขึ้น

4.4 ขั้นตอนการดำเนินการวิจัย

ผู้วิจัยมีวิธีการดำเนินงานจำนวน 6 ขั้นตอน ได้แก่ 1) วิเคราะห์เนื้อหาการทดลองวิทยาศาสตร์ 2) ออกแบบผังงานการใช้การ์ตูนแอนิเมชัน 3) วางโครงเรื่องของการ์ตูนแอนิเมชัน 4) ออกแบบการดำเนินเรื่องของการ์ตูนแอนิเมชัน 5) พัฒนาสื่อมัลติมีเดียแบบมีปฏิสัมพันธ์ และ 6) ทดสอบและประเมินความพึงพอใจของการใช้สื่อมัลติมีเดียแบบมีปฏิสัมพันธ์ แสดงผลดังภาพที่ 1



ภาพที่ 1: ขั้นตอนการดำเนินการวิจัย

จากภาพที่ 1 สามารถอธิบายรายละเอียดในแต่ละขั้นตอนได้ดังนี้

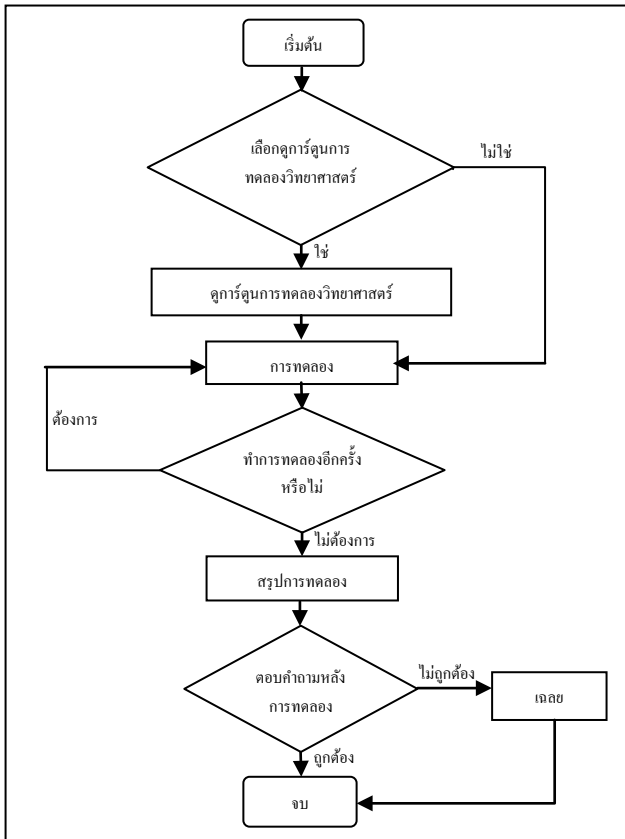
1) วิเคราะห์เนื้อหาการทดลองวิทยาศาสตร์

การวิเคราะห์เนื้อหาการทดลองวิทยาศาสตร์ นั้นจะต้องศึกษาเนื้อหาวิชาวิทยาศาสตร์ของประถมศึกษาปีที่ 2 จำนวน 6 หน่วยการเรียนรู้ จากนั้นคัดเลือกเนื้อหาการทดลอง

วิทยาศาสตร์ที่จะนำมาจัดทำสื่อมัลติมีเดีย ทั้งหมด 4 หน่วยการเรียนรู้ ประกอบด้วย 8 การทดลอง

2) ออกแบบผังงานการใช้การ์ตูนแอนิเมชัน

หลังจากได้นิเื่อหาการทดลองที่จะใช้ทำสื่อมัลติมีเดียเรียบร้อยแล้ว จากนั้นจึงทำการออกแบบผังงานที่จะใช้ในการดำเนินการจัดทำสื่อมัลติมีเดีย ตามภาพที่ 2



ภาพที่ 2: ผังงานของสื่อมัลติมีเดียแบบมีปฏิสัมพันธ์

3) วางโครงเรื่องของการ์ตูนแอนิเมชัน

โครงเรื่องของการ์ตูนแอนิเมชัน ในแต่ละเรื่องจะมีการทดลองแทรกอยู่ภายใน โดยแต่ละเรื่องแบ่งเป็นตอนได้ดังนี้ 1) ของเล่นชิ้นใหม่ของพอล 2) แม่เหล็กดึงดูดผู้เขียนมหัศจรรย์ 3) แม่เหล็กกับแม่เหล็ก 4) หน้าหนาวมาแล้วสินะ 5) แรงไฟฟ้าในสิ่งของ 6) ไฟติดได้ยังไงนะ 7) อะไรเอ่ยนำไฟฟ้า และ 8) เทียบบ้านคุณยาย

4) ออกแบบการดำเนินเรื่องของการ์ตูนแอนิเมชัน

ขั้นตอนของการออกแบบการดำเนินเรื่องของการ์ตูนแอนิเมชันจะมีการนำโครงเรื่องที่ได้วางเอาไว้มาเขียนเป็น Story board โดยมีการออกแบบหน้าต่างของการทดลอง

เสียงพูด และเสียงดนตรีที่ใช้ในการประกอบการ์ตูนแอนิเมชันทั้งหมด

5) พัฒนาสื่อมัลติมีเดีย

พัฒนาโดยใช้โปรแกรม Adobe Flash CS6 และ ภาษา Action Script3.0

6) ทดสอบและประเมินความพึงพอใจของการใช้สื่อมัลติมีเดียแบบมีปฏิสัมพันธ์

การทดสอบสื่อมัลติมีเดียใช้วิธีการทดสอบแบบ Black Box Testing และการประเมินความพึงพอใจของการใช้สื่อมัลติมีเดีย ประเมินความพึงพอใจโดยกลุ่มผู้ใช้งานคือนักเรียนชั้นประถมศึกษาปีที่ 2 โรงเรียนบ้านศรีไค จำนวน 30 คน

5. ผลการดำเนินการวิจัย

5.1 ผลที่ได้จากการพัฒนาสื่อมัลติมีเดียแบบมีปฏิสัมพันธ์

ผู้วิจัยได้พัฒนาสื่อมัลติมีเดียแบบมีปฏิสัมพันธ์เรื่องการทดลองวิทยาศาสตร์บนระบบปฏิบัติการแอนดรอยด์ สำหรับนักเรียนชั้นประถมศึกษาปีที่ 2 ซึ่งประกอบด้วยเนื้อหาการทดลองทั้งหมด เรื่อง โดย 8 มีตัวอย่างการทดลองแสดงเป็นภาพดังนี้



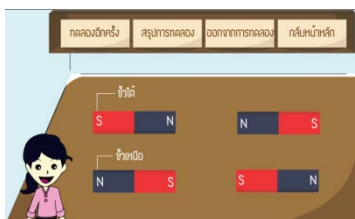
ภาพที่ 3: หน้าแรกของสื่อมัลติมีเดียแบบมีปฏิสัมพันธ์



ภาพที่ 4: การทดลองการดูดซึมน้ำของวัสดุต่างๆ



ภาพที่ 5: การทดลองการดูดแม่เหล็กของวัสดุต่างๆ



ภาพที่ 6: การทดลองการดูดและการผลักของแม่เหล็ก 2 ตัว



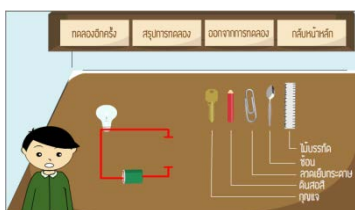
ภาพที่ 7: การทดลองการเกิดไฟฟ้าสถิตในวัตถุต่างๆ



ภาพที่ 8: การทดลองการเกิดแรงไฟฟ้าระหว่างวัตถุ 2 ชนิด



ภาพที่ 9: การทดลองการต่อวงจรไฟฟ้า



ภาพที่ 10: การทดลองฉนวนและตัวนำไฟฟ้า



ภาพที่ 11: การทดลองการอุ้มน้ำของดิน 3 ชนิด(ดินร่วน ดินทรายและดินเหนียว)

5.2 ผลการประเมินประสิทธิภาพสื่อมัลติมีเดีย

ผู้วิจัยได้นำสื่อมัลติมีเดียแบบมีปฏิสัมพันธ์เรื่องการทดลองวิทยาศาสตร์บนระบบปฏิบัติการแอนดรอยด์ สำหรับนักเรียนชั้นประถมศึกษาปีที่ 2 ไปทดลองใช้กับนักเรียนชั้นประถมศึกษาปีที่ 2 โรงเรียนบ้านศรีโค จำนวน 30 คน เพื่อหาประสิทธิภาพของสื่อมัลติมีเดีย ผลปรากฏดังตารางที่ 1

ตารางที่ 1 : ประสิทธิภาพของสื่อมัลติมีเดียแบบมีปฏิสัมพันธ์เรื่องการทดลองวิทยาศาสตร์บนระบบปฏิบัติการแอนดรอยด์ ตามเกณฑ์ 80/80 (E1/E2)

เกณฑ์	ค่าประสิทธิภาพ
E1	81.00
E2	89.75

จากตารางที่ 1 สรุปได้ว่าสื่อมัลติมีเดียแบบมีปฏิสัมพันธ์เรื่องการทดลองวิทยาศาสตร์บนระบบปฏิบัติการแอนดรอยด์ สำหรับนักเรียนชั้นประถมศึกษาปีที่ 2 มีประสิทธิภาพตามเกณฑ์ (81.00/89.75)

5.3 ผลการประเมินความพึงพอใจ

กำหนดระดับความพึงพอใจของการใช้สื่อมัลติมีเดียแบบมีปฏิสัมพันธ์ ดังนี้ [6]

- 4.51 – 5.00 หมายถึง มากที่สุด
- 3.51 – 4.50 หมายถึง มาก
- 2.51 – 3.50 หมายถึง ปานกลาง
- 1.51 – 2.50 หมายถึง น้อย
- 1.00 – 1.50 หมายถึง น้อยที่สุด

ผู้วิจัยได้ศึกษาความพึงพอใจของผู้ใช้สื่อมัลติมีเดียแบบมีปฏิสัมพันธ์ โดยใช้แบบสอบถามความพึงพอใจจากผู้ใช้งานจำนวน 30 คน ผลปรากฏดังตารางที่ 2

ตารางที่ 2 : ผลประเมินความพึงพอใจของผู้ใช้สื่อมัลติมีเดียแบบมีปฏิสัมพันธ์เรื่องการทดลองวิทยาศาสตร์บนระบบปฏิบัติการแอนดรอยด์

รายการประเมิน	\bar{X}	S.D.	การแปลผล
1) การ์ตูนแอนิเมชันทำให้ผู้เรียนเกิดความสนุกสนานเพลิดเพลิน	4.45	0.71	มาก
2) การดำเนินเรื่องของการ์ตูนสอดคล้องกับการทดลอง	4.40	0.49	มาก
3) การทดลองทำให้ผู้เรียนเข้าใจง่าย	4.50	0.56	มาก

รายการประเมิน	\bar{X}	S.D.	การแปลผล
4) ตัวการ์ตูนอธิบายการทดลองได้เข้าใจและผู้เรียนสามารถทำตามได้ง่าย	4.67	0.60	มากที่สุด
5) ตัวการ์ตูนสามารถสรุปการทดลองได้เข้าใจและถูกต้อง	4.60	0.66	มากที่สุด
6) ภาพและตัวการ์ตูนมีความสวยงามและเหมาะสม	4.57	0.67	มากที่สุด
7) สีของปุ่มต่างๆมีความเหมาะสม	4.43	0.76	มาก
8) ตัวอักษรมีขนาดที่เหมาะสมและอ่านง่าย	4.23	2.06	มาก
9) เสียงมีความดังที่เหมาะสมและชัดเจน	4.63	0.48	มากที่สุด
10) เสียงดนตรีประกอบมีความเหมาะสม	4.50	0.85	มาก
เฉลี่ยรวม	4.50	0.75	มาก

จากตารางที่ 2 สรุปได้ว่าความพึงพอใจของผู้ใช้สื่อมัลติมีเดียแบบมีปฏิสัมพันธ์ มีความพึงพอใจอยู่ในระดับมาก ($\bar{X} = 4.50$, $S.D. = 0.75$)

6. สรุปผลการวิจัย

ผลของการวิจัย พบว่า สื่อมัลติมีเดียแบบมีปฏิสัมพันธ์เรื่องการทดลองวิทยาศาสตร์บนระบบปฏิบัติการแอนดรอยด์ สำหรับนักเรียนชั้นประถมศึกษาปีที่ 2 ที่พัฒนาขึ้นมีประสิทธิภาพตามเกณฑ์ 80/80 เมื่อนำสื่อมัลติมีเดียแบบมีปฏิสัมพันธ์ไปทดลองใช้กับนักเรียนชั้นประถมศึกษาปีที่ 2 โรงเรียนบ้านศรีไค จำนวน 30 คน พบว่าผู้เรียนส่วนใหญ่มีความพึงพอใจอยู่ในระดับมาก การวิจัยครั้งนี้ สรุปได้ว่า สื่อมัลติมีเดียแบบมีปฏิสัมพันธ์เรื่องการทดลองวิทยาศาสตร์บนระบบปฏิบัติการแอนดรอยด์ สำหรับนักเรียนชั้นประถมศึกษาปีที่ 2 ที่พัฒนาขึ้น มีประสิทธิภาพและสามารถนำไปใช้เป็นสื่อประกอบในการให้ความรู้หรือใช้ประกอบการเรียนการสอนได้

7. กิตติกรรมประกาศ

ขอขอบคุณ อาจารย์ วิรัตน์ บุญเดิม อาจารย์ประจำชั้นประถมศึกษาปีที่ 2 โรงเรียนบ้านศรีไค ที่ช่วยสนับสนุนเนื้อหาวิชาวิทยาศาสตร์ และเอกสารต่างๆ ในการทำวิจัย

เอกสารอ้างอิง

- [1] สถาบันพัฒนาคุณภาพวิชาการ(พว.).วิทยาศาสตร์ ป.2. กรุงเทพฯ. บริษัทพัฒนาคุณภาพวิชาการ (พว.)จำกัด. 2555.
- [2] การสร้างสื่อการสอนแบบมัลติมีเดีย [ออนไลน์]. สืบค้นจาก:<http://learning.pitlokcenter.com/captivate/train-meaning.htm> [5 กุมภาพันธ์ 57]
- [3] มนตรี แยมกสิกร.เกณฑ์ประสิทธิภาพในงานวิจัย และพัฒนาสื่อการสอน:ความแตกต่าง 90/90 Standard และ E1/E2.วารสารศึกษาศาสตร์ ปีที่ 19 ฉบับที่ 1 เดือนตุลาคม 2550-มกราคม 2551. หน้า 1-16.
- [4] เนรมิต สุดชนะ. การพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนกลุ่มสาระการเรียนรู้วิทยาศาสตร์ เรื่องระบบสุริยะ ชั้นประถมศึกษาปีที่4. [วิทยานิพนธ์ครุศาสตรมหาบัณฑิต].สกลนคร.มหาวิทยาลัยราชภัฏสกลนคร.2551.
- [5] วรพรรณ บุคคีด้วง.พัฒนาบทเรียนคอมพิวเตอร์มัลติมีเดียกลุ่มสาระวิทยาศาสตร์ เรื่อง สารในชีวิตประจำวัน ชั้นมัธยมศึกษาที่ 1. [วิทยานิพนธ์ครุศาสตรมหาบัณฑิต]. สกลนคร. มหาวิทยาลัยราชภัฏสกลนคร.2551.
- [6] ราตรี นันทสุนนธ์. 2554. การวิจัยในชั้นเรียนและการวิจัยพัฒนาการเรียนการสอน(การวิจัยสำหรับครูและบุคลากรทางการศึกษา). หน้า 133 – 136กรุงเทพมหานคร.

ผลการสังเคราะห์แนวทางการประยุกต์ใช้สื่ออีดีแอลทีวีพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ตด้วยระบบพี่เลี้ยง

Practical Guidelines for eDLTV Applications for Learning with Tablet Computer by Using Mentoring Technique

นรากร ศรีวาปี (Narakorn Srivapee)¹ วรปภา อารีราษฎร์ (Worapapha Arreerard)² สุรัตน์ ดวงชาตม (Surat Duangchatom)³

¹นักศึกษาลัทธิศึกษาศาสตร์ สาขาเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏมหาสารคาม

²คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏมหาสารคาม

³ผู้อำนวยการสำนักงานเขตพื้นที่การศึกษาประถมศึกษามหาสารคาม เขต 1

Koi.narakorn@gmail.com, dr.worapapha@hotmail.com, surat.d@hotmail.com

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อ 1) สังเคราะห์องค์ประกอบแนวทางการประยุกต์ใช้สื่ออีดีแอลทีวี พัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ตด้วยระบบพี่เลี้ยง และ 2) สังเคราะห์กิจกรรมการส่งเสริมครูประยุกต์ใช้สื่ออีดีแอลทีวี ส่งเสริมการพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ต โดยใช้ระบบ พี่เลี้ยง กลุ่มเป้าหมายในการวิจัยเป็นผู้เชี่ยวชาญ จำนวน 12 คน จาก 4 กลุ่ม 1) กลุ่มนักเทคโนโลยีสารสนเทศ 2) กลุ่มผู้บริหารโรงเรียนและศึกษานิเทศก์ 3) กลุ่มวิทยากรการประยุกต์ใช้สื่ออีดีแอลทีวีและการพัฒนาแอปพลิเคชัน และ 4) กลุ่มนักวิชาการศึกษาจากมหาวิทยาลัย เครื่องมือที่ใช้ในการวิจัย คือ แบบสอบถามความคิดเห็นของผู้เชี่ยวชาญที่มีต่อองค์ประกอบและแบบสอบถามความคิดเห็นของผู้เชี่ยวชาญที่มีต่อกิจกรรมการส่งเสริมครูประยุกต์ใช้สื่ออีดีแอลทีวีในการวิจัย คือ ค่าเฉลี่ย (\bar{X}) และส่วนเบี่ยงเบนมาตรฐาน (SD)

ผลการวิจัยพบว่า 1) องค์ประกอบแนวทางการประยุกต์ใช้สื่ออีดีแอลทีวีพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ตด้วยระบบพี่เลี้ยง ประกอบด้วย 3 ด้าน คือ 1.1) ความสำคัญของการใช้สื่ออีดีแอลทีวีเพื่อการเรียนรู้ 1.2) การประยุกต์ใช้สื่ออีดีแอลทีวี และ 1.3) หน่วยงานสนับสนุน โดยความคิดเห็นของผู้เชี่ยวชาญที่มีต่อองค์ประกอบของแนวทาง สรุปโดยรวมอยู่ในระดับมากที่สุด ($\bar{X}=4.79$, $SD=0.40$) และ 2) ผลการสังเคราะห์กิจกรรมการส่งเสริมครูประยุกต์ใช้สื่ออีดีแอลทีวี ส่งเสริมการพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ตโดยใช้ระบบพี่เลี้ยง ประกอบด้วย 4 ขั้นตอน คือ 2.1) ขั้นสร้างความเข้าใจ 2.2) ขั้นออกแบบกิจกรรมการเรียนรู้ 2.3) ขั้นพัฒนาแอปพลิเคชัน

และกิจกรรม และ 2.4) ขั้นสรุปและประเมินผล โดยความคิดเห็นของผู้เชี่ยวชาญที่มีต่อกิจกรรมการส่งเสริมครูประยุกต์ใช้ สรุปโดยรวมอยู่ในระดับมากที่สุด ($\bar{X}=4.80$, $SD=0.36$)

คำสำคัญ : สื่ออีดีแอลทีวี, การประยุกต์ใช้สื่อ, ระบบพี่เลี้ยง การจัดการเรียนรู้แบบใช้ปัญหาเป็นฐาน, สื่ออิเล็กทรอนิกส์

Abstract

The research aimed at synthesizing strategies and training activities in eDLTV application for learning with tablet by using a mentoring technique. The target population was twelve participants consisting of 4 major experts in information technology, administrator and educational supervision, eDLTV and application trainers, and academicians. The instrument were 2 questionnaires for expert. The statistics used were mean and standard deviation.

The research showed that three major factors of the practical application of eDLTV for education were 1) importance of ICT for education, 2) ICT application for education, and 3) supporting organization. In addition, the find showed that the average level of the opinion of the experts towards three major factors was very high ($\bar{X}=4.79$, $S.D.=.40$). Regarding eDLTV application for learning with tablet, the research showed that the application of eDLTV for learning with tablet by using a mentoring technique consisted of 4 steps: 1) understanding, 2) design of learning activities, 3) development of

application and activities, and 4) evaluation and conclusion. The finding showed that the average level of the opinion of the experts towards the training activities in eDLTV application for education was very high ($\bar{X} = 4.80$, $S.D. = 0.36$).

Keywords: eDLTV Media, Application, Mentoring

Technique, Problem-Based Learning Activities, Electronic Media

1. บทนำ

โครงการแท็บเล็ตพีซีเพื่อการศึกษาไทย (One Tablet Per Child) ได้ดำเนินการในปี พ.ศ. 2555 โดยการมอบคอมพิวเตอร์แท็บเล็ตให้นักเรียนชั้นประถมศึกษาปีที่ 1 ทั่วประเทศ และในปี พ.ศ. 2556 สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน หรือ สพฐ. ได้ดำเนินการพัฒนาเครื่องมือการสร้างแอปพลิเคชันเพื่อการเรียนรู้ เพื่อให้ครูนำมาเป็นเครื่องมือในการพัฒนาเนื้อหา จัดทำเป็นแอปพลิเคชัน [3] โดยจัดอบรมครูทั่วประเทศ จำนวน 5 รุ่น โดยใช้เครื่องมือการพัฒนาแอปพลิเคชันที่ สพฐ. พัฒนาขึ้น [8] การอบรมครูที่ สพฐ. ดำเนินการมีข้อจำกัด คือ เป็นการอบรมครูจำนวน 1,000 คน แต่ครูที่สังกัด สพฐ. ที่ต้องอบรมนั้น มีจำนวนมาก ดังนั้นสำนักงานเขตพื้นที่การศึกษาต้องดำเนินการขยายผลการอบรมให้ครอบคลุมครูในสังกัด โดยเฉพาะอย่างยิ่งครูผู้สอนชั้นประถมศึกษาปีที่ 2 ที่ต้องใช้คอมพิวเตอร์แท็บเล็ตในการเรียนการสอน ในปีการศึกษา 2556

สำนักงานเขตพื้นที่การศึกษาประถมศึกษามหาสารคาม เขต 1 หรือ สพป.มค. 1 ได้เห็นความสำคัญของการอบรมครู โดยเฉพาะอย่างยิ่งครูผู้สอนชั้นประถมศึกษาปีที่ 2 ตามนโยบายของ สพฐ. [6] และยังได้สนับสนุนให้ครูใช้สื่อไอซีทีหรือพัฒนาสื่อไอซีที เพื่อนำมาเป็นเครื่องมือในการเรียนการสอน โดยให้โรงเรียนในสังกัดใช้สื่ออีดีแอลทีวี หรือ eDLTV ซึ่งเป็นสื่ออิเล็กทรอนิกส์ จากการจัดการเรียนการสอนของ โรงเรียนวังไกลกังวล [2] ที่ครูหรือนักเรียนสามารถเรียนรู้ได้ทุกเวลา ในปีการศึกษา 2556 สพป.มค. 1 มีนโยบายให้ครูประยุกต์ใช้สื่ออีดีแอลทีวี สู่การเรียนการสอน และมีนโยบายอบรมครูชั้นประถมศึกษาปีที่ 2 พัฒนาแอปพลิเคชันเพื่อการเรียนรู้ ตามนโยบายของ สพฐ. [6]

มหาวิทยาลัยราชภัฏมหาสารคามเป็นสถาบันการศึกษาเพื่อพัฒนาท้องถิ่น โดยคณะเทคโนโลยีสารสนเทศ ได้ดำเนินการ

ส่งเสริมการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ของชุมชน ตามนโยบายของรัฐบาล [2] และเป็นหน่วยงานหลักในการเผยแพร่สื่ออีดีแอลทีวี ให้กับโรงเรียนในเขตพื้นที่บริการ จากการดำเนินงานที่ผ่านมา คณะเทคโนโลยีสารสนเทศ มีบุคลากรที่เป็นวิทยากรแกนนำในการอบรมการพัฒนาแอปพลิเคชันร่วมกับ สพฐ. และเป็นวิทยากรแกนนำในการเผยแพร่สื่ออีดีแอลทีวี ให้กับโรงเรียนในเขตพื้นที่บริการ ครอบคลุมจังหวัดมหาสารคาม ขอนแก่น กาฬสินธุ์ และร้อยเอ็ด จากการดำเนินการดังกล่าวส่งผลให้ คณะเทคโนโลยีสารสนเทศ มีบุคลากรเป็นวิทยากรแกนนำ สามารถให้ความรู้ในการพัฒนาแอปพลิเคชันด้วยเครื่องมือของ สพฐ. เป็นที่เลี้ยงให้คำแนะนำการพัฒนาแอปพลิเคชันและการประยุกต์ใช้สื่ออีดีแอลทีวี ให้กับครูในเขตพื้นที่บริการได้

จากผลการอบรมครู ของ สพฐ. ที่ยังไม่ครอบคลุมจำนวนครูได้ทั่วถึง อีกทั้งจากนโยบายของสพป.มค. 1 ในการอบรมครูชั้นประถมศึกษาปีที่ 2 ให้สามารถประยุกต์ใช้สื่ออิเล็กทรอนิกส์จากอีดีแอลทีวีสู่การพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ต โดยอาศัยบุคลากรของคณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏมหาสารคาม เป็นที่เลี้ยงในการอบรม เนื่องจากที่เลี้ยงเป็นผู้เชี่ยวชาญที่มีความรู้ความสามารถ และมีประสบการณ์ในการถ่ายทอด มีทักษะการอบรม [1] เป็นที่ปรึกษา ให้ความช่วยเหลือหรือการให้คำแนะนำ การมีระบบที่เลี้ยงเพื่อให้คำแนะนำในการพัฒนาความรู้ ทักษะความสามารถของครูที่เข้ารับการส่งเสริมหรือการอบรม การพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ต ในระหว่างการอบรมหรือในขั้นตอนการนำสู่การปฏิบัติการสอนที่โรงเรียน จะทำให้ครูได้รับการพัฒนาและได้รับการแนะนำจากผู้มีความรู้และประสบการณ์โดยตรง

จากนโยบายและแนวทางที่กล่าวมา เพื่อให้การพัฒนาครูชั้นประถมศึกษาปีที่ 2 บรรลุผลตามเป้าหมายและวัตถุประสงค์ ผู้วิจัยมีความสนใจ สังเคราะห์แนวทางการประยุกต์ใช้สื่ออีดีแอลทีวีพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ตด้วยระบบที่เลี้ยง เพื่อให้ครูชั้นประถมศึกษาปีที่ 2 ซึ่งเป็นผู้ที่ต้องสอนการใช้แท็บเล็ตตามนโยบายการจัดการเรียนการสอนด้วยคอมพิวเตอร์แท็บเล็ตและนโยบายการใช้สื่ออีดีแอลทีวี มีที่เลี้ยงที่

มีความเชี่ยวชาญให้การแนะนำและเป็นที่ปรึกษา ในการพัฒนาครูของ สพป.มค. 1 ประยุกต์ใช้สื่อ ไอซีทีในการเรียนการสอนต่อไป

2. วัตถุประสงค์

2.1 เพื่อสังเคราะห์องค์ประกอบแนวทางการประยุกต์ใช้สื่อไอทีแอลทีวีพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ตด้วยระบบพี่เลี้ยง

2.2 เพื่อสังเคราะห์กิจกรรมการส่งเสริมครูประยุกต์ใช้สื่อไอทีแอลทีวีสู่การพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ตโดยใช้ระบบพี่เลี้ยง

3. วรรณกรรมที่เกี่ยวข้อง

Becta ICT Research [10] ได้ศึกษาผลการใช้งานแท็บเล็ตพีซีประกอบการเรียนการสอนใน โรงเรียนระดับประถมศึกษาจำนวน 12 แห่ง ในประเทศอังกฤษ ช่วงระหว่างปี ค.ศ.2004-2005 พบว่า การใช้แท็บเล็ตพีซี โดยให้ผู้เรียนและผู้สอนมีแท็บเล็ตพีซีเป็นของตนเองอย่างทั่วถึง เป็นปัจจัยสำคัญที่ช่วยเพิ่มการใช้งานอย่างมีประสิทธิภาพ การใช้แท็บเล็ตพีซีช่วยเพิ่มแรงจูงใจของผู้เรียนและมีผลกระทบในทางบวกต่อผลสัมฤทธิ์ทางการเรียน รวมทั้งสนับสนุนให้เกิดการเรียนรู้ด้วยตนเอง ช่วยส่งเสริมให้เกิดการค้นคว้าและการเข้าถึงองค์ความรู้นอกห้องเรียนอย่างกว้างขวาง รวมทั้งส่งเสริมการเรียนรู้แบบมีส่วนร่วมของผู้เรียน การใช้แท็บเล็ตพีซีช่วยส่งเสริมการใช้เทคโนโลยีในการเรียนการสอนและส่งเสริมให้มีการพัฒนาหลักสูตรหรือการจัดการเรียนการสอนที่มีเทคโนโลยีสารสนเทศเป็นส่วนประกอบมากขึ้น สอดคล้องกับ Gregory Pryor and Vernon Bauer [11] ได้ศึกษาผลการใช้แท็บเล็ตพีซีพบว่า เทคโนโลยีแท็บเล็ตพีซี ทำให้การเรียนการสอนมีชีวิตชีวา สภาพและบรรยากาศของห้องเรียนจะมีการแลกเปลี่ยนและมีการโต้ตอบทั้งผู้สอนและผู้เรียน ด้วยการเชื่อมต่อเครือข่ายไร้สายที่แตกต่างจากการใช้คอมพิวเตอร์ทั่วไป John E. Anderson, Paul H. Schwager and Richard L. Kerns [12] ได้ศึกษาผลการยอมรับและการใช้เทคโนโลยีเพื่อประเมินการยอมรับการใช้แท็บเล็ตพีซี จากคณาจารย์มหาวิทยาลัยในประเทศสหรัฐอเมริกา พบว่า ตัวแปรในการยอมรับการใช้แท็บเล็ตพีซี คือ ความคาดหวังและความสมัคร

ใจ เป็นแรงขับเคลื่อนที่สำคัญ ทั้งนี้ยังพบว่า อายุ เพศ และประสบการณ์มีผลต่อการยอมรับ โดยเฉพาะอย่างยิ่ง เพศหญิงที่มีอายุมากจะให้ความสำคัญในการใช้แท็บเล็ตพีซีค่อนข้างต่ำ

การพัฒนาแอปพลิเคชันเพื่อการเรียนรู้ [7] สพฐ. ได้จัดอบรมครูพัฒนาแอปพลิเคชันเพื่อการเรียนรู้ผู้แท็บเล็ต โดยมีศึกษานิเทศก์เป็นผู้ช่วยในการอบรม ได้จำแนกแอปพลิเคชันเป็น 3 ประเภท คือ 1) แอปพลิเคชันรูปแบบเสริมการเรียนรู้ (Learning Media) ที่นำเสนอเนื้อหาเพื่อให้นักเรียนสามารถเรียนรู้ได้ด้วยตนเอง 2) แอปพลิเคชันรูปแบบเสริมการสอน (Instruction Media) ที่ใช้เป็นตัวช่วยครูในการสอน และ 3) แอปพลิเคชันรูปแบบสร้างองค์ความรู้ (Constructional Media) ที่ช่วยให้ผู้เรียนสามารถสร้างสรรค์ชิ้นงาน ผลการดำเนินงานพบว่า ครูมีทัศนคติเชิงบวกต่อการใช้นวัตกรรมเครื่องมือการพัฒนาแอปพลิเคชันที่ สพฐ. พัฒนาขึ้นในระดับมาก

มหาวิทยาลัยราชภัฏมหาสารคาม [4] ได้นำเครื่องมือของ สพฐ. จากเว็บไซต์ www.otpcappcon.com อบรมครูพัฒนาแอปพลิเคชันบนแท็บเล็ต จำนวน 181 คน ผลการอบรมพบว่า ผู้เข้ารับการอบรมสามารถพัฒนาแอปพลิเคชันที่มีคุณภาพ จำนวน 172 คน คิดเป็น ร้อยละ 95.03 ของผู้เข้ารับการอบรมและผู้เข้ารับการอบรมมีความพึงพอใจในระดับมากที่สุด

สมเจตน์ ภูศรี และคณะ [5] ได้สำรวจการนำสื่อไอทีแอลทีวีไปใช้ของสมาชิกเครือข่าย จำนวน 664 โรงเรียน พบว่า ครูผู้สอนนำสื่อไอทีแอลทีวีไปใช้เป็นแหล่งเรียนรู้เพื่อการศึกษา ค้นคว้าเพิ่มเติม ผู้บริหารและศึกษานิเทศก์ได้ส่งเสริมการนำสื่อไอทีแอลทีวีไปใช้ในห้องสมุด ห้องโสตฯ เผยแพร่ผ่านเว็บไซต์ เพื่อให้ครูและนักเรียนได้ศึกษาค้นคว้า ตลอดจนนำไปใช้เป็นเครื่องมือประกอบการแนะนำและนิเทศ

สุเดือนเพ็ญ กงคะจันทร์ และคณะ [9] กล่าวว่าระบบพี่เลี้ยงเป็นระบบที่ให้ผู้ที่มีความรู้ความสามารถซึ่งเรียกกันว่าพี่เลี้ยง (Mentor) ให้คำปรึกษาแนะนำและช่วยเหลือรุ่นน้องหรือผู้ที่อยู่ในระดับต่ำกว่า (Mentee) ระบบพี่เลี้ยงช่วยให้พนักงานเป็นคนที่เก่งได้อย่างรวดเร็ว โดยมีพี่เลี้ยงเป็นผู้สอนงาน ให้กำลังใจให้การสนับสนุนและให้ความช่วยเหลือให้มีโอกาสได้เติบโตได้รับความก้าวหน้าในหน้าที่การงาน โดยให้โอกาสหรือเวทีที่จะแสดงผลงานแสดงฝีมือ และความสามารถในการทำงาน ในขณะที่ Kram [13] กล่าวว่าระบบพี่เลี้ยงเป็นรูปแบบทางวัฒนธรรมอย่าง

หนึ่งที่สำคัญในองค์กรที่มีประสบการณ์หรือความเชี่ยวชาญมากกว่าให้ความสนใจดูแลบุคคลที่มีประสบการณ์น้อยกว่าและให้การสนับสนุนทางด้านอาชีพและจิตใจรายบุคคล นอกจากนี้

Linda [14] กล่าวว่าพี่เลี้ยงที่มีประสิทธิภาพก็ต้องมีลักษณะดังนี้ 1) เป็นผู้ที่มีประสบการณ์และได้รับการนับถือจากคนในองค์กร 2) แสดงให้เห็นถึงทักษะที่ดีในการพัฒนาผู้อื่น 3) เข้าใจวิธีการที่จะทำให้ผู้ที่อยู่ในความดูแลเรียนรู้ได้ดีที่สุด 4) มีช่องทางของข้อมูลและบุคคลที่สามารถช่วยเหลือในหน้าที่การงานได้ 5) เปิดเผยตรงไปตรงมา 6) สามารถเข้ากันได้กับผู้ที่อยู่ในความดูแล และ 7) มีความผูกพันต่อองค์กร

จากการศึกษารวบรวม ผู้วิจัยได้นำแนวคิดระบบพี่เลี้ยง การประยุกต์ใช้สื่ออีดีแอลทีวี และการพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ตโดยใช้เครื่องมือของ สพฐ. เป็นแนวทางการสังเคราะห์องค์ประกอบแนวทางการประยุกต์ใช้สื่ออีดีแอลทีวีพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ตด้วยระบบพี่เลี้ยง ที่เป็นองค์ประกอบนำมาสู่การสังเคราะห์กิจกรรมการอบรมครู ซึ่งจะส่งผลให้ครูสามารถประยุกต์ใช้สื่ออีดีแอลทีวีผู้การพัฒนาแอปพลิเคชัน นำสู่การเรียนการสอนในชั้นเรียนได้อย่างมีประสิทธิภาพและมีประสิทธิผลสอดคล้องกับนโยบายการจัดการเรียนการสอนด้วยคอมพิวเตอร์แท็บเล็ตและนโยบายการพัฒนาครูใช้สื่อไอซีทีของ สพฐ. และ สพป.มค. 1 ต่อไป

4. วิธีการดำเนินงาน

4.1 กลุ่มเป้าหมาย

การวิจัยครั้งนี้ เป็นผู้เชี่ยวชาญที่สำเร็จการศึกษาระดับปริญญาเอก และมีประสบการณ์ในการประยุกต์ใช้สื่ออีดีแอลทีวีและการพัฒนาแอปพลิเคชัน จำนวน 12 คน โดยแบ่งเป็น 4 กลุ่ม กลุ่มละ 4 คน คือ 1) กลุ่มนักเทคโนโลยีสารสนเทศ 2) กลุ่มผู้บริหารโรงเรียนและศึกษานิเทศก์ 3) กลุ่มวิทยากรการประยุกต์ใช้สื่ออีดีแอลทีวีและการพัฒนาแอปพลิเคชัน และ 4) กลุ่มนักวิชาการศึกษาจากมหาวิทยาลัย

4.2 เครื่องมือที่ใช้ในการวิจัย

แบบสอบถามความคิดเห็นผู้เชี่ยวชาญ จำนวน 2 ชุด คือแบบสอบถามความคิดเห็นที่มีต่อแนวทางและแบบสอบถามที่มีต่อกิจกรรมการประยุกต์ใช้สื่อ

4.3 ขั้นตอนการดำเนินงาน

4.3.1 การศึกษาเอกสาร บทความ และงานวิจัยเกี่ยวกับหลักการ แนวคิด และทฤษฎีที่เกี่ยวข้อง

4.3.2 สังเคราะห์องค์ประกอบแนวทางการประยุกต์ใช้สื่ออีดีแอลทีวีผู้การพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ตด้วยระบบพี่เลี้ยง และ สังเคราะห์กิจกรรมการส่งเสริมครูประยุกต์ใช้สื่ออีดีแอลทีวีผู้การพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ตโดยใช้ระบบพี่เลี้ยง

4.3.3 จัดประชุมกลุ่มย่อยโดยผู้เชี่ยวชาญจำนวน 12 คน เพื่อศึกษาความคิดเห็นที่มีต่อองค์ประกอบแนวทางการประยุกต์ใช้สื่ออีดีแอลทีวีและกิจกรรมการส่งเสริมครู

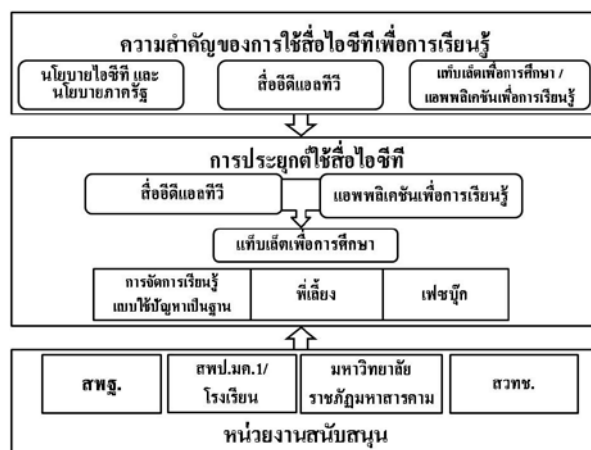
4.3.4 วิเคราะห์ข้อมูลและสรุปผลการวิจัย

4.4 การสังเคราะห์ข้อมูล

ผู้วิจัยนำแบบสำรวจที่ได้จากกลุ่มเป้าหมายมาวิเคราะห์โดยใช้สถิติพื้นฐาน ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน

5. ผลการดำเนินงาน

5.1 ผลการสังเคราะห์องค์ประกอบแนวทางการประยุกต์ใช้สื่ออีดีแอลทีวี พัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ตด้วยระบบพี่เลี้ยง พบว่าแนวทางมีองค์ประกอบ 3 ด้าน ดังแสดงในภาพที่ 1



ภาพที่ 1 : องค์ประกอบแนวทางการประยุกต์ใช้สื่ออีดีแอลทีวี

จากภาพที่ 1 องค์ประกอบแนวทางการประยุกต์ใช้สื่ออีดีแอลทีวีพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ตด้วยระบบพี่เลี้ยง ประกอบด้วยองค์ประกอบ 3 ด้าน ดังนี้

5.1.1 ด้านความสำคัญของการใช้สื่อไอซีทีเพื่อการเรียนรู้ โดยในการดำเนินงานส่งเสริมครูจะได้นำนโยบายและ

เป้าหมายในส่วนนี้ มาเป็นกรอบในการดำเนินการ ประกอบด้วย 1) นโยบายไอซีทีและนโยบายภาครัฐ 2) สื่ออีดีแอลทีวี และ 3) แท็บเล็ตเพื่อการศึกษา/แอปพลิเคชันเพื่อการเรียนรู้

5.1.2 ด้านการประยุกต์ใช้สื่อไอซีที โดยในการดำเนินงานส่งเสริมครูจะดำเนินการจัดกิจกรรมส่งเสริมครู ที่เริ่มจากสื่ออีดีแอลทีวี และแอปพลิเคชันเพื่อการเรียนรู้ โดยการศึกษาและเรียนรู้ เพื่อนำสู่การพัฒนาแอปพลิเคชันในรูปแบบต่าง ๆ ตามความต้องการ เพื่อนำมาเป็นสื่อตามนโยบายแท็บเล็ตเพื่อการศึกษา ซึ่งจะต้องมีเครื่องมือเป็นองค์ประกอบ คือ การจัดการเรียนรู้แบบใช้ปัญหาเป็นฐาน พี่เลี้ยง และเฟซบุ๊ก

5.1.3 ด้านหน่วยงานสนับสนุน ประกอบด้วย สพฐ. สพป.มค.1 มหาวิทยาลัยราชภัฏมหาสารคาม และ สวทช. โดย สพฐ. สนับสนุนเครื่องมือการพัฒนาแอปพลิเคชัน สพป.มค.1 สนับสนุนงบประมาณ บุคลากรที่เป็นศึกษานิเทศก์เพื่อการตรวจผลงานที่ครูผลิต มหาวิทยาลัยราชภัฏสนับสนุนเครื่องมือวิทยากรและพี่เลี้ยง และ สวทช. สนับสนุนสื่ออีดีแอลทีวี

5.2 ผลการสังเคราะห์กิจกรรมการส่งเสริมครูประยุกต์ใช้สื่ออีดีแอลทีวีสู่การพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ตด้วยระบบพี่เลี้ยง

ผู้วิจัยสังเคราะห์กิจกรรม ซึ่งเป็นขั้นตอนสำหรับการส่งเสริมครูประยุกต์ใช้สื่ออีดีแอลทีวีสู่การพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ตด้วยระบบพี่เลี้ยง โดยนำหลักการและแนวคิดจากองค์ประกอบที่ได้จากการสังเคราะห์ตามภาพที่ 1 ได้ผลการสังเคราะห์ แสดงในภาพที่ 2

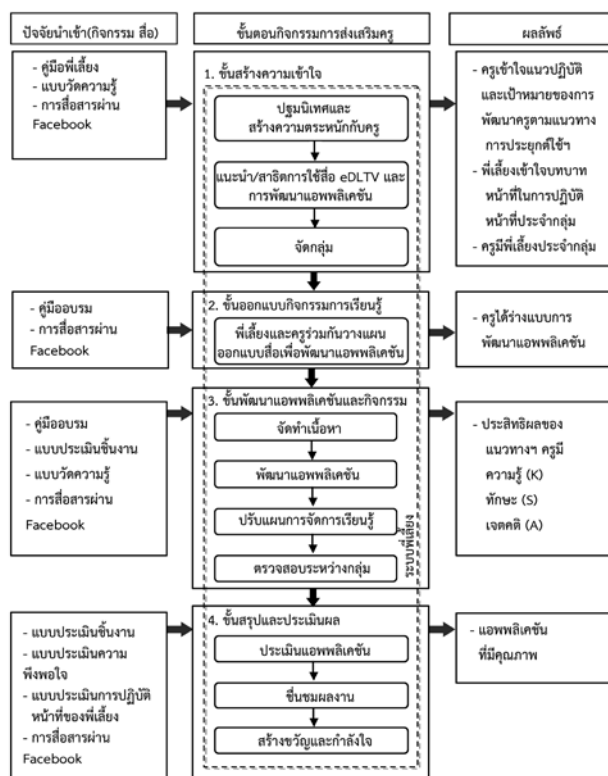
จากภาพที่ 2 มีกิจกรรมการส่งเสริมครูประยุกต์ใช้สื่อ ประกอบด้วย 4 ชั้น ระยะเวลา 3 วัน จำนวน 18 ชั่วโมง ดังนี้

5.2.1 ชั้นสร้างความเข้าใจ ระยะเวลา 3 ชั่วโมง เป็นขั้นตอนการสร้างความเข้าใจกับครูกลุ่มเป้าหมายโดยการปฐมนิเทศและสร้างความตระหนักให้กับครู กำหนดบทบาทหน้าที่ของพี่เลี้ยง และกิจกรรมที่ต้องดำเนินการร่วมกัน

5.2.2 ชั้นออกแบบกิจกรรมการเรียนรู้ ระยะเวลา 3 ชั่วโมง เป็นขั้นตอนที่ให้พี่เลี้ยงและครูร่วมกัน วางแผนออกแบบกิจกรรมการเรียนรู้เพื่อนำไปใช้ในขั้นตอนต่อไป

5.2.3 ชั้นพัฒนาแอปพลิเคชันและกิจกรรม ระยะเวลา 9 ชั่วโมง เป็นขั้นตอนการอบรมปฏิบัติการเพื่อพัฒนาครูตามแนวทางการประยุกต์ใช้สื่ออีดีแอลทีวี

5.2.4 ชั้นสรุปและประเมินผลกิจกรรม ระยะเวลา 3 ชั่วโมง เป็นขั้นตอนการประเมินแอปพลิเคชันที่ครูพัฒนาขึ้น นำเสนอในระดับกลุ่มและนำเสนอในระดับภาพรวมเพื่อการสะท้อนผลการพัฒนาครูและผลงานที่พัฒนาขึ้น



ภาพที่ 2 : กิจกรรมการส่งเสริมครูประยุกต์ใช้สื่อ

5.3 ผลการศึกษาความคิดเห็นผู้เชี่ยวชาญที่มีต่อองค์ประกอบและกิจกรรมการส่งเสริมครูประยุกต์ใช้สื่ออีดีแอลทีวีของผู้เชี่ยวชาญ แสดงในตารางที่ 1

ตารางที่ 1 : ผลการศึกษาความคิดเห็นของผู้เชี่ยวชาญ

รายการ	\bar{X}	S.D.	การแปลผล
1) ความเหมาะสมขององค์ประกอบแนวทางการประยุกต์ใช้สื่ออีดีแอลทีวี	4.79	0.40	มากที่สุด
ด้านที่ 1 ความสำคัญของการใช้สื่อไอซีทีเพื่อการเรียนรู้	4.87	0.34	มากที่สุด
ด้านที่ 2 การประยุกต์ใช้	4.73	0.44	มากที่สุด
ด้านที่ 3 หน่วยงานสนับสนุน	4.77	0.43	มากที่สุด

2) ความเหมาะสมของกิจกรรมการส่งเสริมครูประยุกต์ใช้สื่ออีดีแอลทีวี	4.80	0.36	มากที่สุด
ขั้น 1 สร้างความเข้าใจ	4.93	0.20	มากที่สุด
ขั้น 2 ออกแบบกิจกรรมการเรียนรู้	4.67	0.49	มากที่สุด
ขั้น 3 พัฒนาแอปพลิเคชันและกิจกรรม	4.85	0.29	มากที่สุด
ขั้น 4 สรุปและประเมินผลกิจกรรม	4.76	0.44	มากที่สุด

จากตารางที่ 1 พบว่า ความคิดเห็นของผู้เชี่ยวชาญที่มีต่อองค์ประกอบแนวทางการประยุกต์ใช้สื่ออีดีแอลทีวีในภาพรวมอยู่ในระดับมากที่สุด ($\bar{x} = 4.79$, S.D. = 0.40) โดยองค์ประกอบทั้ง 3 ด้านมีความเหมาะสมในระดับมากที่สุด และความคิดเห็นของผู้เชี่ยวชาญที่มีต่อกิจกรรมการส่งเสริมครูประยุกต์ใช้สื่อในภาพรวมอยู่ในระดับมากที่สุด ($\bar{x} = 4.80$, S.D. = 0.36) โดยกิจกรรมทั้ง 4 ขั้นตอนมีความเหมาะสมในระดับมากที่สุด

6. สรุปผล

จากการวิจัยโดยการสังเคราะห์เอกสารและงานวิจัยที่เกี่ยวข้องและจากผลการศึกษาความคิดเห็นของผู้เชี่ยวชาญที่มีต่อแนวทางการประยุกต์ใช้สื่ออีดีแอลทีวีพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนคอมพิวเตอร์แท็บเล็ตด้วยระบบพีดีเอช พบว่าองค์ประกอบของแนวทางมี 3 ด้าน ได้แก่ 1) ด้านความสำคัญของการใช้สื่อไอซีทีเพื่อการเรียนรู้ 2) ด้านการประยุกต์ใช้สื่อไอซีที และ 3) ด้านหน่วยงานสนับสนุน และในส่วนกิจกรรมการส่งเสริมครูประยุกต์ใช้สื่อ มี 4 ขั้นตอน ได้แก่ 1) ขั้นสร้างความเข้าใจ 2) ขั้นออกแบบกิจกรรมการเรียนรู้ 3) ขั้นพัฒนาแอปพลิเคชันและกิจกรรม และ 4) ขั้นสรุปและประเมินผลกิจกรรม โดยผู้เชี่ยวชาญมีความคิดเห็นต่อผลการสังเคราะห์ในครั้งนี้อยู่ในระดับมากที่สุด ทั้งนี้ ผู้วิจัยจะนำองค์ประกอบและขั้นตอนของกิจกรรมการส่งเสริมครูประยุกต์ใช้สื่ออีดีแอลทีวีที่สังเคราะห์ได้นำไปสู่การดำเนินงานวิจัยในระยะต่อไป

เอกสารอ้างอิง

[1] กิติมา ปรีดีศิลป์, “การบริหารและการนิเทศการศึกษาเบื้องต้น,” กรุงเทพฯ : อักษรการ, 2532.
 [2] โครงการเทคโนโลยีสารสนเทศเพื่อการศึกษาของโรงเรียนในชนบท, จาก <http://edltv.thai.net/index.php?mod=Message&op=aboutus>

สืบค้นเมื่อ 10 เมษายน 2556.
 [3] โครงการประกวดสื่อสร้างสื่อการเรียนรู้สู่แท็บเล็ต, จาก <http://www.otpcappcon.com> สืบค้นเมื่อ 5 กุมภาพันธ์ 2556,
 [4] มหาวิทยาลัยราชภัฏมหาสารคาม, “รายงานผลการอบรม โครงการศูนย์ทางไกลฯ มหาวิทยาลัยราชภัฏมหาสารคาม,” มหาสารคาม : คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏมหาสารคาม, 2556.
 [5] สมเจตน์ ภูศรี และคณะ, “รูปแบบการเผยแพร่ถ่ายทอดและพัฒนาสื่อ eDLTV เครื่องข่ายความร่วมมือมหาวิทยาลัยราชภัฏมหาสารคาม,” มหาสารคาม : มหาวิทยาลัยราชภัฏมหาสารคาม, 2554.
 [6] สำนักงานเขตพื้นที่การศึกษาประถมศึกษามหาสารคาม เขต 1, “แผนปฏิบัติการประจำปีงบประมาณ 2556,” มหาสารคาม : สำนักงานเขตพื้นที่การศึกษาประถมศึกษามหาสารคาม เขต 1, 2556.
 [7] สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน, “การศึกษาผลการอบรมครูโดยใช้นวัตกรรมระบบพัฒนาสื่อการเรียนการสอนตามโครงการประกวดสร้างสื่อการเรียนรู้...สู่แท็บเล็ต,” กรุงเทพฯ : สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน, 2556.
 [8] สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน, “คู่มืออบรมสร้างสื่อการเรียนรู้สู่แท็บเล็ต,” กรุงเทพฯ : สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน, 2556.
 [9] สุเดือนเพ็ญ คงคะจันทร์ และคณะ, “เอกสารประกอบการอบรมหลักสูตรนักรับบริหารงานส่งเสริมสุขภาพระดับกลาง รุ่นที่ 6/2550 ระหว่างวันที่ 24 สิงหาคม – 3 กันยายน 2550 ณ โรงแรมเซ็นทรัลเพลส จังหวัดสมุทรสาคร,” 2550.
 [10] Becta ICT Research, “Tablet PCs in schools : Case study report,” [online] 2004. Available from: http://www.oro.open.ac.uk/6407/1/BTE_case_study_print.pdf [2013, August 9].
 [11] Gregory Pryor and Vernon Bauer, “Building a better biology lab? Testing tablet PC technology in a core laboratory course,” *Journal of College Science Teaching*, November/December, pp. 44-48, 2008.
 [12] John E. Anderson Paul H. Schwager and Richard L. Kerns, “The drivers for acceptance of tablet PCs by faculty in a college of business,” *Journal of information system education*, vol. 17, No.4, pp. 429-437, 2006.
 [13] Kram, K.E., “Mentoring at Work : Developmental Relationships in Organizational Life,” *Glenview, Illinois:Scott Foresman and Company*, 1985.
 [14] Linda, A.H. “Becoming A Manager,” *Boston, Massachusetts: Harvard Business School Press*, 1992.

รูปแบบการส่งเสริมการเรียนรู้โรงเรียนทสรข. ด้วยกิจกรรมค่ายอาสาเยาวชนทั่วถิ่นไทยเรียนรู้
 ได้ใต้ร่มพระบารมีเครือข่ายมหาวิทยาลัยราชภัฏมหาสารคาม
**A Model of Camping Activities for Learning Improvement through
 RajabhatMahaSarakham University eDLTV Networks of Schools Participating in
 the Information Technology Project under the Initiatives of Her Royal Highness
 Princess Maha Chakri Sirindhorn**

ธวัชชัย สหพงษ์ (ThawatchaiSahapong)¹ วรปภา อารีราษฎร์ (WorapaphaArreerard)²

ธรัช อารีราษฎร์ (TharachArreerard)³ และ ธเนตร ยืนสุข (ThanateYuensook)⁴

คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏมหาสารคาม

¹tone302@hotmail.com, ²dr.worapapha@hotmail.com, dr.tharach@hotmail.com

บทคัดย่อ

การวิจัยมีวัตถุประสงค์เพื่อ 1) ศึกษารูปแบบการจัดกิจกรรมค่ายอาสาที่มหาวิทยาลัยจัดร่วมกับโรงเรียน ทสรข. และ 2) ศึกษาความคิดเห็นของผู้เข้าร่วมกิจกรรมค่ายอาสาตามรูปแบบที่พัฒนาขึ้น กลุ่มเป้าหมายในการวิจัยเป็นผู้บริหาร ครูและนักเรียนโรงเรียน ทสรข. ในภาคตะวันออกเฉียงเหนือ จำนวน 10 แห่ง รวมทั้ง 1,011 คน ประกอบด้วย ผู้บริหารและครู 65 คน และนักเรียน 946 คน เครื่องมือที่ใช้ในการวิจัย คือ รูปแบบการจัดกิจกรรมค่ายอาสา โรงเรียน ทสรข. และ แบบสอบถามความพึงพอใจผู้เข้าร่วมกิจกรรม

ผลการวิจัยพบว่า 1) รูปแบบกิจกรรมค่ายอาสาที่มหาวิทยาลัยจัดร่วมกับโรงเรียน ทสรข. เป็นรูปแบบที่แสดงถึงความสัมพันธ์ของ 5 องค์ประกอบ คือ ผู้เข้าร่วมกิจกรรม สื่อการเรียนรู้ โรงเรียน ทสรข. หน่วยงานสนับสนุน ตัวชี้วัด และกิจกรรมค่ายอาสา จำนวน 4 กลุ่ม คือ กลุ่ม A กิจกรรมระบบและเครือข่ายคอมพิวเตอร์ กลุ่ม B กิจกรรมการประยุกต์ใช้สื่ออิเล็กทรอนิกส์เพื่อการเรียนการสอน กลุ่ม C กิจกรรมการส่งเสริมครูและนักเรียนพัฒนาหนังสืออิเล็กทรอนิกส์ และ กลุ่ม D กิจกรรมการส่งเสริมการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ ผลการประเมินความคิดเห็นของผู้เชี่ยวชาญ จำนวน 12 คน ที่มีต่อรูปแบบกิจกรรมค่ายอาสาโรงเรียน ทสรข. พบว่ารูปแบบที่พัฒนาขึ้นมีความเหมาะสมในระดับมากที่สุด (ค่าเฉลี่ย 4.67 และส่วนเบี่ยงเบนมาตรฐาน 0.49) และ 2) ผลการศึกษาความคิดเห็นของผู้เข้าร่วมกิจกรรมค่ายอาสาตามรูปแบบที่พัฒนาขึ้น จาก 5 ตัวชี้วัด พบว่า ผู้เข้าร่วมกิจกรรมค่ายอาสาที่มีความพึงพอใจต่อกิจกรรมเฉลี่ยโดยรวมอยู่ในระดับมากที่สุด (ค่าเฉลี่ย 4.52 และส่วนเบี่ยงเบนมาตรฐาน 0.59)

คำสำคัญ: สื่ออิเล็กทรอนิกส์ สื่ออิเล็กทรอนิกส์ โรงเรียน ทสรข.

กิจกรรมค่ายอาสา นวัตกรรม

Abstract

The objective of the research were to 1) study a model of camping activities for learning improvement through Rajabhat Maha Sarakham University eDLTV Networks of Schools Participating in the Information Technology Project under the Initiatives of Her Royal Highness Princess Maha Chakri Sirindhorn and 2) study the opinions of participants towards camping activities for learning improvement. The target population was one hundred and eleven participants consisting of 65 of administrators and teachers and 946 students from ten Schools participating in the Information Technology Project under the Initiatives of Her Royal Highness Princess Maha Chakri Sirindhorn in the northeast of Thailand. The instrument were a model of camping activities for learning improvement and questionnaire.

The research finding a model of camping activities related to five major factors consisting of trainees, teaching materials, schools participating in the Information Technology Project under the Initiatives of Her Royal Highness Princess Maha Chakri Sirindhorn, supporting organization, indicators and camping activities. The camping activities were composed of four groups: a) activity of computer system and network, b) activity of practical application of eDLTV for learning, c) activity of development the electronic book for teachers and students, and d) activity of using

tablet computer for learning. The studying showed that the average opinion of twelve experts towards the appropriateness of the model of camping activities was very high ($\bar{X}=4.67$, $S.D. = 0.49$). In addition, the research showed that the average level of the satisfaction of trainees with five indicators camping activities was very high ($\bar{X}=4.52$, $S.D. = 0.59$).

Keywords: Electronic Media, eDLTV Media, Schools Participating in the Information Technology Project under the Initiatives of Her Royal Highness Princess Maha Chakri Sirindhorn, Camping Activities, Innovation

1. บทนำ

โรงเรียน ทสรช. [1] เป็นโรงเรียนในโครงการเทคโนโลยีสารสนเทศตามพระราชดำริสมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี มีจำนวน 85 แห่ง ในภาคตะวันออกเฉียงเหนือ มีโรงเรียนในโครงการ ทสรช. 10 แห่ง มหาวิทยาลัยราชภัฏมหาสารคาม [2] เป็นสมาชิกเครือข่าย เผยแพร่ ถ่ายทอดและพัฒนาสื่อการเรียนการสอนบนระบบ e-Learning (eDL-Square) ร่วมกับสำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ หรือ สวทช. โดยศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ หรือ เนคเทค (National Electronics and Computer Technology Center: NECTEC) เพื่อให้มหาวิทยาลัยเป็นศูนย์กลางในการส่งเสริมการเผยแพร่ ถ่ายทอดและพัฒนาสื่ออีดีแอลทีวี (Electronic Distance Learning Television : eDLTV) ในการเรียนการสอนให้กับโรงเรียนที่เข้าร่วมโครงการ ส่งเสริมให้โรงเรียนนำระบบอีดีแอลทีวี ไปใช้ในการพัฒนาคุณภาพการศึกษาของนักเรียน และเพื่อให้มีการวิจัยพัฒนาต่อยอด ปรับปรุงกระบวนการเรียนการสอน โดยการใช้ประโยชน์จากระบบ eDL-Square ตลอดจนให้คำปรึกษาแนะนำแก่โรงเรียนในโครงการ ทสรช. ใช้ระบบอีดีแอลทีวีในการเรียนการสอน ภายใต้การส่งเสริม สนับสนุน การจัดกิจกรรม และแลกเปลี่ยนประสบการณ์การใช้อีดีแอลทีวี โดย สวทช.

สื่ออีดีแอลทีวี [3] เป็นสื่อที่มูลนิธิการศึกษาทางไกลผ่านดาวเทียมและโครงการเทคโนโลยีสารสนเทศตามพระราชดำริสมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี ได้จัดทำขึ้นเพื่อร่วมเทิดพระเกียรติพระบาทสมเด็จพระเจ้าอยู่หัวฯ เนื่องในโอกาสมหามงคลเฉลิมพระชนมพรรษา 80 พรรษา 5 ธันวาคม 2550 โดยการนำเนื้อหาที่ถ่ายทอดออกอากาศทางสถานีโทรทัศน์

การศึกษาทางไกลผ่านดาวเทียม หรือ DLTV จากโรงเรียนวังไกลกังวล อำเภอหัวหิน จังหวัดประจวบคีรีขันธ์ ในระดับมัธยมศึกษา ระดับประถมศึกษา และระดับอาชีวศึกษาหรือกลุ่มการทำงานอาชีพ ที่ประกอบด้วย แผนการสอน วิดีทัศน์การสอน สไลด์ ใบงาน ใบความรู้ และแบบทดสอบ รวบรวมและจัดเก็บด้วยระบบบริหารจัดการอีดีแอลสแควร์ (eDL-Square) เพื่อให้โรงเรียนในโครงการ ทสรช. สามารถนำมาใช้ในการจัดการเรียนการสอนเพื่อการเรียนรู้แบบไม่ผ่านเครือข่าย หรืออีเลิร์นนิ่งแบบออฟไลน์ (Off-Line e-Learning) การเรียนการสอนแบบเครือข่ายภายในโรงเรียนด้วยระบบแลน (Local Area Network: LAN) และการเรียนการสอนแบบเครื่องเดียว (Stand Alone) [10] มหาวิทยาลัยราชภัฏมหาสารคาม ได้เข้าร่วมเป็นสมาชิกเครือข่าย เผยแพร่ ถ่ายทอด และพัฒนาสื่อการเรียนการสอนบนระบบ e-Learning หรือสื่ออีดีแอลทีวี ตั้งแต่ พ.ศ. 2552 ในปี พ.ศ. 2555 มหาวิทยาลัยได้มุ่งเน้นการเผยแพร่สื่ออีดีแอลทีวีพร้อมกับการจัดกิจกรรมค่ายอาสาร่วมกับโรงเรียนเครือข่าย โรงเรียน ทสรช. ในภาคตะวันออกเฉียงเหนือ 10 เป็นโรงเรียนเครือข่ายของมหาวิทยาลัย เพื่อให้การดำเนินงานบรรลุตามเป้าหมายและสอดคล้องกับนโยบายในการดำเนินงานของโครงการ ในการส่งเสริมการใช้สื่ออีดีแอลทีวีให้กับโรงเรียน ทสรช. ผู้วิจัยจึงสนใจศึกษารูปแบบการส่งเสริมการเรียนรู้โรงเรียน ทสรช. ด้วยกิจกรรม “ค่ายอาสาเยาวชนทั่วถิ่นไทยเรียนรู้ได้ได้ร่วมพระบารมี” เครือข่าย มหาวิทยาลัยราชภัฏมหาสารคาม ซึ่งผู้วิจัยหวังว่ากิจกรรมในรูปแบบดังกล่าวจะเป็นกิจกรรมที่ส่งเสริมให้สมาชิกที่เข้าร่วมมีความพึงพอใจในการเข้าร่วมกิจกรรม จากการจัดกิจกรรมในรูปแบบต่างๆ เพื่อสร้างความรู้ความเข้าใจในการใช้สื่ออีดีแอลทีวี สามารถบูรณาการและประยุกต์ใช้สื่ออีดีแอลทีวี ร่วมกับสื่อไอซีทีให้เป็นประโยชน์ต่อการเรียนการสอน ต่อไป

2. วัตถุประสงค์การวิจัย

- 2.1 เพื่อศึกษารูปแบบการจัดกิจกรรมค่ายอาสาที่มหาวิทยาลัยจัดร่วมกับโรงเรียน ทสรช.
- 2.2 เพื่อศึกษาความพึงพอใจของผู้เข้าร่วมกิจกรรมค่ายอาสาตามรูปแบบที่พัฒนาขึ้น

3. วรรณกรรมที่เกี่ยวข้อง

โรงเรียน ทสรช.[1] ส่วนใหญ่เป็นโรงเรียนที่อยู่ในชนบทห่างไกล มีปัญหาด้านการจัดการเรียนการสอน เช่น ครูไม่ครบชั้น ครูสอนไม่ตรงวุฒิการศึกษา และจำนวนครูไม่สอดคล้องกับชั้นเรียนและกลุ่มสาระการเรียนรู้ที่สอน จึงได้นำการเรียนการสอนทางไกลผ่านดาวเทียม หรือ DLTV ที่ถ่ายทอดออกอากาศมาใช้ในการเรียนการสอน แต่ก็ยังประสบปัญหาคือ นักเรียนจดไม่ทันการสอนของครู โรงเรียนวังไกลกังวล และครูยังต้องการสื่อการเรียนการสอนจากระบบ DLTV เพื่อให้นักเรียนได้ศึกษาเรียนรู้

กิจกรรมค่ายอาสา [6] เป็นกิจกรรมที่จัดโดยมหาวิทยาลัยร่วมกับโรงเรียนเครือข่ายที่ขอรับสื่ออีดีแอลทีวี จัดกิจกรรม 1 วัน ด้วยกิจกรรมหลายรูปแบบ ที่มุ่งเน้นส่งเสริมการเรียนรู้สื่ออีดีแอลทีวีให้ครูโรงเรียนเครือข่ายได้เรียนรู้และเข้าใจ สามารถประยุกต์ใช้สื่อเพื่อการเรียนการสอนได้อย่างมีประสิทธิภาพ

รูปแบบกิจกรรมค่ายอาสาอีดีแอลทีวีมหาวิทยาลัยราชภัฏมหาสารคาม [4] เป็นกิจกรรมความร่วมมือ ระหว่างมหาวิทยาลัยราชภัฏมหาสารคาม สำนักงานเขตพื้นที่การศึกษา โรงเรียน และชุมชน ประกอบด้วย 3 กิจกรรม คือ 1) กิจกรรมการเรียนรู้สื่ออีดีแอลทีวีกับการพัฒนาหนังสืออิเล็กทรอนิกส์ 2) กิจกรรมการเรียนรู้สื่ออีดีแอลทีวีกับการใช้คอมพิวเตอร์แท็บเล็ต และ 3) กิจกรรมการเรียนรู้สื่ออีดีแอลทีวีกับฮาร์ดแวร์ ระยะเวลาในการจัดกิจกรรม 1 วัน ผลการศึกษาความพึงพอใจของผู้เข้าร่วมกิจกรรมค่ายอาสาตามรูปแบบที่พัฒนาขึ้นอยู่ในระดับมากที่สุด และ ผลการศึกษาความคิดเห็นของผู้เข้าร่วมกิจกรรมค่ายอาสา พบว่า 1)กิจกรรมค่ายอาสา มีประโยชน์ 96.05% 2) ประโยชน์ของกิจการเรียนรู้สื่ออีดีแอลทีวีกับการพัฒนาหนังสืออิเล็กทรอนิกส์ 90.40% 3) ประโยชน์ของกิจการเรียนรู้สื่ออีดีแอลทีวีกับการใช้แท็บเล็ต 89.83% 4) ประโยชน์ของกิจการเรียนรู้สื่ออีดีแอลทีวีกับฮาร์ดแวร์ 87.57 % 5) กิจกรรมค่ายอาสาทำให้ครูมีความสนใจในการใช้สื่ออีดีแอลทีวีมากขึ้น 87.01 % 6) กิจกรรมค่ายอาสาทำให้ครู นักเรียน และผู้บริหารได้เรียนรู้ร่วมกัน 86.44 % 7) กิจกรรมค่ายอาสาเชื่อมความสัมพันธ์โรงเรียนกับชุมชน 84.75 % และ 8) กิจกรรมค่ายอาสาให้ความรู้ด้านการบูรณาการสื่อไอซีทีเพื่อการเรียนการสอน 84.18 %

โครงการ RMU-eDL เครือข่ายความร่วมมือมหาวิทยาลัยราชภัฏมหาสารคามกับโรงเรียน ทสรช. ในภาคตะวันออกเฉียงเหนือ [9] ผู้วิจัยได้พัฒนารูปแบบการพัฒนาศักยภาพครูและบุคลากร

โรงเรียน ทสรช. ประกอบด้วย 5 กิจกรรม คือ 1) การชี้แจงแนวทาง 2) การลงนามความร่วมมือ 3) การพัฒนาสื่ออิเล็กทรอนิกส์ 4) การผลิตรายงานการวิจัย และ 5) การให้ความช่วยเหลือตรวจสอบ และประเมิน ผลการดำเนินการตามรูปแบบพบว่าโรงเรียนศึกษาสงเคราะห์นางรอง จังหวัดบุรีรัมย์ได้พัฒนาบุคลากร จำนวน 60 คน พัฒนาหนังสืออิเล็กทรอนิกส์ที่มีคุณภาพ 45 เล่ม หลังการจัดกิจกรรมตามรูปแบบพบว่า บุคลากรที่เข้าร่วมกิจกรรมมีความรู้ความเข้าใจในการพัฒนาหนังสืออิเล็กทรอนิกส์ การเขียนรายงานการวิจัย เพิ่มขึ้นอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และมีความพึงพอใจต่อการอบรมเฉลี่ยโดยรวมอยู่ในระดับมากที่สุด

4. วิธีดำเนินการวิจัย

4.1 ขอบเขตของการวิจัย การวิจัยในครั้งนี้เป็นการศึกษารูปแบบการจัดกิจกรรมค่ายอาสาที่มหาวิทยาลัยราชภัฏมหาสารคามจัดร่วมกับโรงเรียน ทสรช. ภายใต้โครงการ “เยาวชนทั่วถิ่นไทยเรียนรู้ได้พร้อมพระบารมี” (เกิดพระเกียรติ 60 พระชันษา มหาจักรีสิรินธร) เพื่อกระตุ้นให้ ผู้บริหาร ครู และนักเรียน ประยุกต์ใช้สื่ออีดีแอลทีวีให้เกิดประโยชน์สูงสุด

4.2 กลุ่มเป้าหมายในการวิจัย เป็นสมาชิกเครือข่ายอีดีแอลทีวีกลุ่มโรงเรียน ทสรช. ภาคตะวันออกเฉียงเหนือ 10 โรงเรียน รวมทั้งหมด 1,011 คน ประกอบด้วย กลุ่มผู้บริหารและครู 65 คน และกลุ่มนักเรียน 946 คน

4.3 เครื่องมือที่ใช้ในการวิจัย คือ รูปแบบการจัดกิจกรรมค่ายอาสา และ แบบสอบถามความพึงพอใจผู้เข้าร่วมกิจกรรม

4.4 ระยะเวลาดำเนินการ เดือน กันยายน – ธันวาคม 2555

4.5 ขั้นตอนการดำเนินการวิจัย ใช้ขั้นตอน PAOR จำแนกเป็น 4 ระยะ ดังนี้

ระยะที่ 1 วางแผน (Plan : P) เพื่อสร้างเครือข่ายความร่วมมือ โดยการประชุมผู้บริหารและครู กลุ่มโรงเรียน ทสรช. โรงเรียนละ 2 คน เข้าร่วมประชุมวางแผนการดำเนินงาน และลงนามบันทึกข้อตกลงความร่วมมือ

ระยะที่ 2 ดำเนินการ (Action : A) เพื่อสังเคราะห์รูปแบบการจัดกิจกรรมค่ายอาสาโรงเรียน ทสรช. ให้พร้อมกับการดำเนิน ดังนี้

1) ประชุมผู้แทนจากโรงเรียน ทสรช. โรงเรียนละ 1 คน รวมทั้งสิ้น 10 คน เพื่อศึกษาแนวทางการจัดกิจกรรมค่ายอาสา

2) สังเคราะห์รูปแบบกิจกรรมคำขอาสา พร้อมจัดทำ เครื่องมือการประเมินผล และ คู่มือกิจกรรม

3) ประชุมกลุ่มย่อย ผู้ทรงคุณวุฒิ จำนวน 12 คน จาก 4 กลุ่ม ๆ ละ 3 คน คือกลุ่มผู้บริหาร กลุ่มครูจากโรงเรียน ทสรช. กลุ่มนักวิชาการศึกษา และกลุ่มศึกษานิเทศก์ เพื่อสรุปผลและ ประเมินรูปแบบกิจกรรมคำขอาสา

ระยะที่ 3 สังเกตการปฏิบัติการ (Operation Observe : O) เพื่อ ศึกษาผลการนำรูปแบบไปใช้ในการจัดกิจกรรมคำขอาสา ดังนี้

1) ประชุมเพื่อวางแผนกำหนด หน้าที่ในการจัดทำ กิจกรรมคำขอาสา

2) มอบหมายหน้าที่และกิจกรรม

3) จัดกิจกรรมคำขอาสาที่โรงเรียนตามรูปแบบกำหนด

ระยะที่ 4 สะท้อนผลกิจกรรม (Activity Reflect : R) เพื่อให้ ผู้เข้าร่วมกิจกรรม ร่วมสรุปโดยเขียนรายงาน เรื่องเล่า ดังนี้

1) หลังจากสิ้นสุดการจัดกิจกรรมคำขอาสา มอบหมาย ให้ครูและนักเรียนเขียนเรื่องเล่า จากการจัดกิจกรรมของครูที่ โรงเรียนและกำหนดวันนำเสนอเรื่องเล่า

2) คณะกรรมการคัดเลือกเรื่องเล่าจากการนำเสนอ โดยมี ผู้แทนจากโรงเรียน ศึกษานิเทศก์ และผู้บริหาร ร่วมคัดเลือก

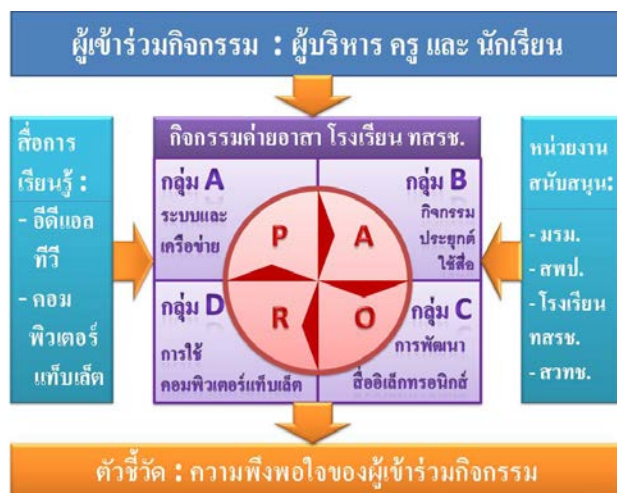
3) สังเคราะห์และสรุป ความรู้ ประโยชน์และผลที่ได้ จากการจัดกิจกรรมคำขอาสาจากเรื่องเล่า

4) นำเสนอรายงานต่อที่ประชุมครูและผู้บริหาร โรงเรียนละ 2 คน โดยเป็นครู 1 คน และผู้บริหาร 1 คน

5) สรุปผลการดำเนินงานและเผยแพร่กิจกรรมคำขอาสา และเรื่องเล่า ที่เว็บไซต์ www.itrmu.net/edlitv

5. ผลการดำเนินการวิจัย

5.1 ผลการศึกษารูปแบบการจัดกิจกรรมคำขอาสาโรงเรียน ทสรช. พบว่า รูปแบบกิจกรรมคำขอาสาเป็นรูปแบบความร่วมมือ ระหว่างที่มหาวิทยาลัยราชภัฏมหาสารคาม กับโรงเรียน ทสรช. สำนักงานเขตพื้นที่การศึกษาประถมศึกษาร้อยเอ็ด เขต 1 และ สำนักงาน สวทช. ในรูปแบบกิจกรรมคำขอาสา “เยาวชนทั่วถิ่น ไทย เรียนรู้ได้ร่วมพระบารมี” (เทิดพระเกียรติ 60 พระชันษา มหาจักรีสิรินธร) ดังภาพที่ 1



ภาพที่ 1: รูปแบบกิจกรรมคำขอาสาโรงเรียน ทสรช.

จากภาพที่ 1 รูปแบบกิจกรรมคำขอาสา โรงเรียน ทสรช. เป็นรูปแบบที่แสดงถึงความสัมพันธ์ของ 5 องค์ประกอบ คือ ผู้เข้าร่วมกิจกรรม สื่อการเรียนรู้ โรงเรียน ทสรช. หน่วยงาน สนับสนุน ตัวชี้วัด และกิจกรรมคำขอาสา จำนวน 4 กลุ่ม รวม ทั้งหมด 12 กิจกรรม ผลการประเมินความคิดเห็นของผู้เชี่ยวชาญ จำนวน 12 คน ที่มีต่อรูปแบบกิจกรรมคำขอาสา โรงเรียน ทสรช. มีความเหมาะสมในระดับมากที่สุด (ค่าเฉลี่ย 4.67 และส่วน เบี่ยงเบนมาตรฐาน 0.49)

5.2 ผลการศึกษาความคิดเห็นของผู้เข้าร่วมกิจกรรมคำข อาสาโรงเรียน ทสรช. ตามรูปแบบที่พัฒนาขึ้น

ผู้วิจัยได้ดำเนินการจัดกิจกรรมคำขอาสาที่โรงเรียนศึกษา สงเคราะห์วัชรบุรี ในวันที่ 21 กันยายน 2555 เป็นกิจกรรมที่ แสดงถึงความสัมพันธ์ ของ 5 องค์ประกอบ ดังนี้



ภาพที่ 2: บุคลากรที่เข้าร่วมกิจกรรมคำขอาสา

1) ผู้เข้าร่วมกิจกรรม ครั้งนี้รวมทั้งสิ้น 1,011 คน ประกอบด้วย กลุ่มผู้บริหารและครู 65 คน และกลุ่มนักเรียน 946 คน

2) สื่อการเรียนรู้ คือ สื่ออีดีแอลทีวี และคอมพิวเตอร์แท็บเล็ต

3) โรงเรียน ทสรช. ที่เข้าร่วม มาจากโรงเรียน ทสรช. ในภาค ตะวันออกเฉียงเหนือ 10 แห่ง ดังนี้

- 3.1) โรงเรียนศึกษาสงเคราะห์รัชชบุรี
- 3.2) โรงเรียนศึกษาสงเคราะห์อำนาจเจริญ
- 3.3) โรงเรียนราชประชานุเคราะห์ 28
- 3.4) โรงเรียนราชประชานุเคราะห์ 29
- 3.5) โรงเรียนราชประชานุเคราะห์ 32
- 3.6) โรงเรียนโสตศึกษาจังหวัดร้อยเอ็ด
- 3.7) โรงเรียนศึกษาสงเคราะห์ขอนแก่น
- 3.8) โรงเรียนศึกษาสงเคราะห์สกลนคร
- 3.9) โรงเรียนราชประชานุเคราะห์ 37
- 3.10) โรงเรียนราชประชานุเคราะห์ 23
- 4) หน่วยงานสนับสนุน 4 หน่วยงาน คือ

4.1) สำนักงานเขตพื้นที่การศึกษาประถมศึกษาร้อยเอ็ด เขต 1 โดย ดร. สุรัตน์ ดวงชาตม ผู้อำนวยการ สพป.รอ.เขต 1 เป็นประธานเปิดกิจกรรมค่ายอาสา

4.2) โรงเรียน ทสรช. เป็นผู้เข้าร่วมกิจกรรม ประกอบด้วย ผู้บริหาร ครู และนักเรียน ทั้งนี้กิจกรรมได้รับความอนุเคราะห์จาก ดร. อนงค์ พิษสิงห์ ผู้อำนวยการโรงเรียนศึกษาสงเคราะห์รัชชบุรี กล่าวรายงานการจัดกิจกรรมค่ายอาสาต่อประธาน

4.3) มหาวิทยาลัยราชภัฏมหาสารคาม โดยทีมงานวิทยากร ผู้ช่วยศาสตราจารย์ ดร. วรปภา อารีราษฎร์ และบุคลากรจากคณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏมหาสารคาม

4.4) สวทช. สนับสนุนสื่ออีดีแอลทีวี ให้กับโรงเรียน ทสรช. และมหาวิทยาลัยราชภัฏมหาสารคาม

5) ตัวชี้วัดกิจกรรมค่ายอาสา 5 ตัวชี้วัด จากการจัดกิจกรรม 4 กลุ่ม 12 กิจกรรม ดังนี้

กลุ่ม A กิจกรรมระบบและเครือข่ายคอมพิวเตอร์ ประกอบด้วย กิจกรรมย่อย 3 กิจกรรม คือ 1) การอบรมให้ความรู้การซ่อมบำรุงคอมพิวเตอร์สำหรับครู 2) การอบรมติดตั้งโปรแกรมคอมพิวเตอร์สำหรับครู และ 3) การอบรมการดูแลระบบและเครือข่ายให้นักเรียน



ภาพที่ 3: กิจกรรมกลุ่ม A ระบบและเครือข่ายคอมพิวเตอร์

กิจกรรมกลุ่ม B การประยุกต์ใช้สื่ออีดีแอลทีวีเพื่อการเรียนการสอน ประกอบด้วย 5 กิจกรรม คือ กิจกรรมการแนะนำการ



ภาพที่ 4: กิจกรรมกลุ่ม B การประยุกต์ใช้สื่ออีดีแอลทีวีใช้สื่ออีดีแอลทีวีเพื่อการเรียนรู้ โดยจัดกิจกรรมแนะนำตามกลุ่มที่เข้าร่วมกิจกรรม 5 กลุ่ม คือ 1) กลุ่มผู้บริหารนำเสนอในภาพรวม 2) กลุ่มครูระดับอนุบาล 3) กลุ่มครูระดับประถมศึกษา 4) กลุ่มครูระดับมัธยมศึกษา และ 5) กลุ่มครูอาชีวศึกษาหรือการงานอาชีพ

กลุ่ม C กิจกรรมส่งเสริมครูและนักเรียนพัฒนาหนังสืออิเล็กทรอนิกส์ ประกอบด้วย 2 กิจกรรม คือ 1) การอบรมครูพัฒนาหนังสืออิเล็กทรอนิกส์เพื่อการเรียนรู้ และ 2) การอบรมนักเรียนพัฒนาสื่อแบบมัลติมีเดียและสื่อปฏิสัมพันธ์



ภาพที่ 5: กิจกรรมกลุ่ม C ส่งเสริมการพัฒนาหนังสืออิเล็กทรอนิกส์

กลุ่ม D กิจกรรมส่งเสริมการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ ประกอบด้วย 2 กิจกรรม คือ 1) กิจกรรมส่งเสริมการใช้คอมพิวเตอร์สำหรับครู และ 2) กิจกรรมการผลิตสื่อวีดิทัศน์ด้วยคอมพิวเตอร์แท็บเล็ตสำหรับนักเรียน



ภาพที่ 6: กิจกรรมกลุ่ม D ส่งเสริมการใช้คอมพิวเตอร์แท็บเล็ต

ผู้วิจัยได้สำรวจตัวชี้วัด โดยใช้แบบสอบถามความพึงพอใจของกลุ่มตัวอย่างที่เข้าร่วมกิจกรรมค่ายอาสา แสดงดังตารางที่ 1

ตารางที่ 1: ความพึงพอใจที่มีต่อกิจกรรมค่ายอาสา

รายการตัวชี้วัดกิจกรรมค่ายอาสา	ค่าเฉลี่ย	S.D.	ความหมาย
1. ความเหมาะสมของเนื้อหา/สาระที่ได้รับจากการจัดกิจกรรมค่ายอาสา	4.51	0.55	มากที่สุด
2. ความเหมาะสมด้านระยะเวลาในการจัดกิจกรรมค่ายอาสา	4.34	0.81	มาก
3. ความรู้และความเข้าใจการประยุกต์ใช้สื่ออีดีแอลทีวีในการจัดการสอนของครูและการพัฒนาการเรียนรู้ของนักเรียน	4.61	0.49	มากที่สุด
4. ความรู้และทักษะในการประยุกต์ใช้สื่ออีดีแอลทีวีให้นักเรียนศึกษาเรียนรู้ในเวลาและนอกเวลาได้อย่างมีประสิทธิภาพ	4.64	0.49	มากที่สุด
5. การเพิ่มความรู้และทักษะทางคอมพิวเตอร์ให้กับครูและนักเรียน	4.50	0.55	มากที่สุด
เฉลี่ยโดยรวม	4.52	0.59	มากที่สุด

จากตารางที่ 1 พบว่า ผู้เข้าร่วมกิจกรรมค่ายอาสาโรงเรียนทสรข. มีความพึงพอใจต่อกิจกรรมค่ายอาสาเฉลี่ยโดยรวมระดับมากที่สุด (ค่าเฉลี่ย 4.52 และส่วนเบี่ยงเบนมาตรฐาน 0.59) เมื่อพิจารณารายการตัวชี้วัดกิจกรรมค่ายอาสาพบว่า มี 4 ตัวชี้วัดอยู่ในระดับมากที่สุด และ 1 ตัวชี้วัดอยู่ในระดับมาก

6. สรุปผลการวิจัย

ผลการศึกษารูปแบบการจัดกิจกรรมค่ายอาสาโรงเรียนทสรข. พบว่า รูปแบบกิจกรรมค่ายอาสาเป็นรูปแบบความร่วมมือมี 5 องค์ประกอบ คือ ผู้เข้าร่วมกิจกรรม สื่อการเรียนรู้ โรงเรียนทสรข. หน่วยงานสนับสนุน ตัวชี้วัด และกิจกรรมค่ายอาสา จำนวน 4 กลุ่ม รวมทั้งหมด 12 กิจกรรม ผลการประเมินความคิดเห็นของผู้เชี่ยวชาญที่มีต่อรูปแบบกิจกรรมค่ายอาสา มีความเหมาะสมในระดับมากที่สุด จากการจัดกิจกรรมค่ายอาสาระยะเวลา 1 วัน ผลการศึกษาความพึงพอใจของผู้เข้าร่วมกิจกรรมค่ายอาสาตามรูปแบบที่พัฒนาขึ้น พบว่าผู้เข้าร่วมกิจกรรมมีความพึงพอใจต่อกิจกรรมในระดับมากที่สุด สอดคล้องการดำเนินงานรูปแบบกิจกรรมค่ายอาสา โรงเรียนเครือข่ายอีดีแอลทีวีมหาวิทยาลัยราชภัฏมหาสารคาม [4] จากการจัดกิจกรรม 26 ครั้ง จำนวน 70 โรงเรียน [6] พบว่าผู้เข้าร่วมกิจกรรมมีความพึงพอใจและกิจกรรมค่ายอาสาเป็นกิจกรรมที่มีประโยชน์ ทำให้ครูและ

นักเรียน เรียนรู้ร่วมกัน และสามารถประยุกต์ใช้สื่ออีดีแอลทีวีในการเรียนการสอนได้อย่างมีประสิทธิภาพ สอดคล้องกับวัตถุประสงค์และเป้าหมายของโครงการ [7],[8]

7. กิตติกรรมประกาศ

การวิจัยครั้งนี้ดำเนินการบรรลุลตามวัตถุประสงค์ ผู้วิจัยขอขอบคุณ โครงการเทคโนโลยีสารสนเทศตามพระราชดำริสมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ หรือ สวทช. ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ หรือเนคเทค (NECTEC) มูลนิธิการศึกษาทางไกลผ่านดาวเทียม สำนักงานเขตพื้นที่การศึกษาประถมศึกษาร้อยเอ็ด เขต 1 โรงเรียนโครงการ ทสรข. ภาคตะวันออกเฉียงเหนือและ บุคลากรคณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏมหาสารคาม

เอกสารอ้างอิง

- [1] โรงเรียน ทสรข. โครงการเทคโนโลยีสารสนเทศตามพระราชดำริ สมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี. เข้าถึงได้จาก <http://www.princess-it.com/> [Online].
- [2] มหาวิทยาลัยราชภัฏมหาสารคาม. รายงานผลการดำเนินงานการส่งเสริมเผยแพร่และพัฒนาเครือข่ายสื่อ eDLTV. มหาวิทยาลัยราชภัฏมหาสารคาม. 2554.
- [3] มหาวิทยาลัยราชภัฏมหาสารคาม. รายงานผลการดำเนินงานการส่งเสริมเผยแพร่และพัฒนาเครือข่ายสื่อ eDLTV. มหาวิทยาลัยราชภัฏมหาสารคาม. 2555.
- [4] วรภา อภิราษฎร์ และ ธรัช อภิราษฎร์. รูปแบบกิจกรรมค่ายอาสาอีดีแอลทีวีมหาวิทยาลัยราชภัฏมหาสารคาม. งานประชุมวิชาการ NCCIT-2013. คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ. 2555.
- [5] สมเจตน์ ภูศรี และคณะ. การส่งเสริมการใช้เทคโนโลยีสารสนเทศและการสื่อสารสู่ท้องถิ่นภายใต้บริบท มหาวิทยาลัยราชภัฏมหาสารคาม. งานประชุมวิชาการ ICSSS 2011. มหาวิทยาลัยราชภัฏมหาสารคาม. 2554.
- [6] คณะเทคโนโลยีสารสนเทศ. รายงานสรุปการจัดกิจกรรมค่ายอาสา 2553-2555. มหาวิทยาลัยราชภัฏมหาสารคาม. 2555.
- [7] สำนักงาน สวทช. รายงานผลการดำเนินงานโครงการเครือข่ายเผยแพร่ถ่ายทอด และพัฒนาสื่อการเรียนการสอนบนระบบ e-Learning (eDL-Square). สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ, กระทรวงวิทยาศาสตร์และเทคโนโลยี. กรุงเทพมหานคร. 2553.

- [8] สำนักงาน สวทช. บันทึกข้อตกลงความร่วมมือ เครือข่ายเผยแพร่ถ่ายทอด และพัฒนาสื่อการเรียนการสอนบนระบบe-Learning (eDL-Square) ระยะที่ 2. สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ, กระทรวงวิทยาศาสตร์และเทคโนโลยี, กรุงเทพมหานคร. 2555.
- [9] ชรัช อารีราษฎร์ และคณะ. รูปแบบการพัฒนาศักยภาพครูและบุคลากร ในการพัฒนาสื่ออิเล็กทรอนิกส์ โครงการ RMU-eDL เครือข่ายความร่วมมือมหาวิทยาลัยราชภัฏมหาสารคามกับโรงเรียน ทสรช. ในภาคตะวันออกเฉียงเหนือ. งานประชุมวิชาการ ICSSS 2013. มหาวิทยาลัยราชภัฏมหาสารคาม. 2556.
- [10] โครงการจัดทำเนื้อหาบนระบบ e-Learning ของการศึกษาทางไกลผ่านดาวเทียมเฉลิมพระเกียรติ เนื่องในโอกาสมหามงคลเฉลิมพระชนมพรรษา 80 พรรษา 5 ธันวาคม 2550. เข้าถึงได้จาก <http://www.edltv.thai.net>. [Online].

เสริมพลังชุมชนด้วยไอที เพื่อการเกษตรแบบพอเพียงที่ยั่งยืน

Empower the Community with IT for the Sustainable Sufficient Agriculture.

นิลubl ทองชัย (Nilubon Tongchai)

สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏกาญจนบุรี

nilubont@yahoo.com

บทคัดย่อ

งานวิจัยนี้นำเสนอเรื่อง “เสริมพลังชุมชนด้วยไอที เพื่อการเกษตรแบบพอเพียงที่ยั่งยืน” มีวัตถุประสงค์ในการพัฒนาเว็บไซต์สำหรับใช้เป็นแหล่งข้อมูลชุมชนแบบออนไลน์ ที่สอดคล้องกับความต้องการของผู้ใช้ซึ่งเป็นสมาชิกในชุมชน โดยรวบรวมข้อมูล ออกแบบและพัฒนาเว็บไซต์ โดยนำเสนอข้อมูลแบบจัดตามกลุ่มเนื้อหา แบ่งได้เป็น 6 ส่วน ใช้ชื่อเว็บไซต์ว่า www.thamakham.com จากนั้นนำไปทดสอบกับกลุ่มทดลองจำนวน 92 คน ก่อนนำไปใช้ถ่ายทอดความรู้การใช้งานเว็บไซต์ 6 ครั้ง มีผู้เข้าร่วม 145 คน ซึ่งผลการประเมินในภาพรวมของ 4 ประเด็นหลักอยู่ที่ระดับมาก นอกจากนี้ผลจากการถ่ายทอดความรู้ ช่วยให้สมาชิกในชุมชนท่ามะขาม มองเห็นความสำคัญของสิ่งที่ชุมชนมี และเป็นอยู่ รวมถึงหันกลับมาทบทวนถึงสิ่งที่ดำเนินอยู่ หรือนำเสนออยู่ในเว็บไซต์ว่าควรจะพัฒนาต่อไปในทิศทางใด ซึ่งสิ่งนี้เองเป็นความร่วมมือที่จะนำไปสู่การพัฒนาฐานข้อมูลชุมชนแบบพอเพียงที่ยั่งยืน

คำสำคัญ: ฐานข้อมูลชุมชน เว็บไซต์ชุมชน

Abstract

This research presented "Empower the Community with IT for the Sustainable Sufficient Agriculture" aiming at developing website to be used as an online-community data source. In order to do that, community information are acquired and represented by means of the content-based structure concept into 6 sections under the www.thamakham.com URL. The website is tested and assessed by 92 participants before producing 6 times of knowledge transferred to other 145 participants. The results show that participants from 6 groups agreed on 4

main issues at a "much level". In addition, the knowledge transferred let the community members look back for who they are? What they have? And how they can achieved together before exchanging knowledge via the provided website? This somehow led the community database to the way of sustained sufficiently development.

Keyword: Community database, Community website

1. บทนำ

จากคำเปรียบที่ว่า เกษตรกรคือกระดูกสันหลังของชาติ บ่งชี้ให้เห็นถึงความสำคัญของอาชีพเกษตรกรที่มีมาตั้งแต่อดีตจนถึงปัจจุบัน แต่เป็นที่น่าสังเกตว่าผู้ประกอบการอาชีพนี้ ส่วนใหญ่มีฐานะค่อนข้างยากจน ทั้งๆ ที่มีอาชีพเกี่ยวข้องกับปากท้องของประชากรทั้งประเทศ ซึ่งถือได้ว่าเป็นความจำเป็นพื้นฐานของการดำรงชีพ และจากการศึกษาพบว่าเกษตรกรส่วนใหญ่ทำการเกษตรและขายผลผลิตตามแบบที่สืบทอดมาแต่โบราณจากบรรพบุรุษ มีเพียงส่วนน้อยที่นำความรู้และนวัตกรรมทางด้านต่างๆ มาช่วยเพื่อเพิ่มผลผลิต หรือบางส่วนนำข้อมูลที่ไร้คุณภาพมาใช้งาน ทำให้ไม่สามารถตอบโจทย์หรือความต้องการของเกษตรกรหรือผู้ใช้ข้อมูลในชุมชนได้

ดังนั้นคำถามต่อมา คือ ทำอย่างไร เกษตรกรเหล่านี้จึงเข้าถึงข้อมูลเชิงลึกเกี่ยวกับสิ่งที่ตัวเองกำลังสนใจหรือดำเนินการอยู่ เพื่อขอคำปรึกษาจากผู้เชี่ยวชาญโดยตรง เพิ่มความเข้มแข็งขึ้นในชุมชน นอกจากนั้น ถ้าหน่วยงานภายนอกต้องการสนับสนุนด้านงบประมาณ หรืองานวิจัยในเรื่องต่างๆ ที่เกี่ยวข้องกับชุมชน สามารถติดต่อโดยตรงได้ที่ใคร นอกเหนือจากข้อมูลรายงานหมู่บ้านชนบทไทย จากข้อมูลพื้นฐานระดับหมู่บ้าน

(กชข 2 ค.) [1] ที่ทางภาครัฐจัดทำไว้ในระดับหมู่บ้าน ซึ่งในช่วงก่อนหน้าสามารถเข้าถึงหรือดาวน์โหลดข้อมูลได้เพียงบางส่วน รวมถึงเอกสารบางไฟล์ไม่สามารถเปิด หรือเปิดได้แต่อ่านไม่ได้ เป็นต้น ซึ่งอาจส่งผลให้ชุมชนเสียโอกาสในการได้รับความร่วมมือจากหน่วยงานภายนอก

เพื่อตอบคำถามดังกล่าว ผู้วิจัยได้มีแนวคิดพัฒนาเว็บไซต์สำหรับใช้เป็นแหล่งข้อมูลชุมชนแบบออนไลน์ ที่สอดคล้องกับความต้องการของผู้ใช้ซึ่งเป็นสมาชิกในชุมชน โดยได้รวบรวมข้อมูล ใช้หลักการออกแบบและพัฒนาเว็บไซต์ นำเสนอข้อมูลโดยจัดตามกลุ่มเนื้อหา สำหรับจัดการและเผยแพร่ข้อมูลตามความต้องการของชุมชน ช่วยให้สมาชิกในชุมชน โดยเฉพาะผู้ประกอบการอาชีพด้านการเกษตรสามารถนำความรู้ที่ได้เพื่อใช้วางแผนการจัดการเพิ่มรายได้จากผลผลิตที่มีอยู่ในปัจจุบัน ให้เกิดประโยชน์อย่างคุ้มค่าที่สุด รวมไปถึงการถ่ายทอดเทคโนโลยีที่ได้ศึกษาให้กับประชากรในชุมชน เพื่อให้เกิดการแลกเปลี่ยนเรียนรู้ที่ยั่งยืนต่อไป [2]

2. วรรณกรรมที่เกี่ยวข้อง

แนวคิดที่เกี่ยวกับการจัดทำเว็บไซต์เพื่อจัดการข้อมูลแบบออนไลน์ สำหรับใช้จัดการและเผยแพร่ข้อมูลตามความต้องการของชุมชน เพื่อการเกษตรแบบพอเพียงที่ยั่งยืน ประกอบด้วย การจัดทำเว็บไซต์ฐานข้อมูลชุมชนออนไลน์ และเกษตรแบบพอเพียงที่ยั่งยืน ซึ่งมีรายละเอียดดังต่อไปนี้

2.1 การจัดทำเว็บไซต์ฐานข้อมูลชุมชน

การจัดทำฐานข้อมูลชุมชน หมายถึง การสำรวจ รวบรวมข้อมูล และวิเคราะห์แยกแยะจำแนกเป็นหมวดหมู่ จัดเก็บไว้ในรูปแบบที่เข้าใจได้ง่ายและเรียกใช้ได้อย่างรวดเร็ว การสำรวจและรวบรวมข้อมูลชุมชนทั่วไปนั้น อาจจัดทำแบบเชิงปริมาณและเชิงคุณภาพ [3]

ในการออกแบบเว็บไซต์ สิ่งสำคัญที่ต้องคำนึงถึงในอันดับต้นๆ คือ การกำหนดเป้าหมายของเว็บไซต์ กำหนดกลุ่มผู้ใช้เรียนรู้ผู้ใช้ เพื่อช่วยให้เราสามารถออกแบบเนื้อหาและการใช้งานเว็บไซต์ได้อย่างเหมาะสม ตรงกับความต้องการของผู้ใช้ อย่างแท้จริง การวางโครงสร้างเว็บไซต์ (Site structure design) เพื่อเป็นแผนผังลำดับการนำเสนอเนื้อหาทั้งหมดบน

เว็บไซต์ให้มองเห็นอย่างเป็นรูปธรรม ช่วยให้ผู้ใช้ไม่สับสนและค้นหาข้อมูลที่ต้องการ ได้อย่างรวดเร็ว วิธีการจัดโครงสร้างเว็บไซต์ที่นิยมใช้ได้แก่ จัดตามกลุ่มเนื้อหา (Content-based structure) และจัดตามกลุ่มผู้ชม (User-based structure)

นอกจากนี้เว็บไซต์ที่มีประสิทธิภาพควรมีคุณลักษณะดังต่อไปนี้ คือ ความเรียบง่าย (Simplicity) ความเป็นเอกลักษณ์ (Identity) เนื้อหาที่มีประโยชน์ (Useful content) ระบบการนำทางใช้งานง่าย (User-friendly navigation) และการใช้งานได้อย่างไม่จำกัด (Compatibility) [4]

มีงานวิจัยที่เกี่ยวข้องกับการจัดทำเว็บไซต์และฐานข้อมูลออนไลน์ไว้หลากหลาย อาทิจัดทำระบบฐานข้อมูลบนเว็บเพื่อนำเสนอข่าวชายแดนภาคใต้ [4] การพัฒนาระบบฐานข้อมูลแพทย์พื้นบ้าน [5] นอกจากนี้ในประเด็นของการจัดเก็บข้อมูลท้องถิ่นควรมีข้อมูลทั่วไปรวมทั้งสิ้น 12 ด้าน และการได้มาซึ่งข้อมูลอาจได้จากการสอบถามโดยตรง สำรองและเก็บข้อมูลท้องถิ่น ศึกษาจากเอกสาร และศึกษาจากเว็บไซต์ [6]

จากงานวิจัย [7] พบว่าชายและหญิงมีความแตกต่างกันอย่างมีนัยสำคัญในด้านพฤติกรรมการใช้เทคโนโลยีสารสนเทศ แต่ไม่มีความแตกต่างอย่างมีนัยสำคัญทางสถิติ ในด้านความพึงพอใจและโอกาสในการเข้าถึงความต้องการข้อมูลข่าวสารจากเทคโนโลยีสารสนเทศ ตลอดจนปัญหาและอุปสรรคที่เกิดขึ้นจากการใช้งาน

2.2 เกษตรแบบพอเพียงที่ยั่งยืน

เศรษฐกิจพอเพียงเป็นแนวพระราชดำริที่พระบาทสมเด็จพระเจ้าอยู่หัวภูมิพลอดุลยเดชฯ ได้ทรงพระราชทานแก่ปวงชนชาวไทย มีหลักการสำคัญ คือ ความพอประมาณ ความมีเหตุผล รวมถึงมีระบบภูมิคุ้มกันในตัวที่ดี ต่อการมีผลกระทบอันเกิดจากการเปลี่ยนแปลงทั้งภายนอกและภายใน [8]

รูปแบบการทำเกษตรกรรมยั่งยืนแบ่งเป็นประเภทต่างๆ คือ ระบบไร่หมุนเวียน ระบบเกษตรผสมผสาน ระบบไร่นาสวนผสม ระบบไร่นาป่าผสมหรือวนเกษตร เกษตรธรรมชาติ เกษตรทฤษฎีใหม่ เกษตรกรรมประณีต และเกษตรอินทรีย์ [9]

จากโครงการวิจัยเอกสารด้านเศรษฐกิจพอเพียง [10] ทำให้ทราบถึงแนวคิดเกี่ยวกับเศรษฐกิจพอเพียง การแก้ปัญหาความยากจน การวิจัยเพื่อการพัฒนาท้องถิ่นหรือชุมชน การเรียนรู้

ของชุมชน ตลอดจนปัจจัยเงื่อนไขต่างๆ ที่ทำให้ชุมชนต้นแบบสามารถปรับใช้แนวคิดเศรษฐกิจพอเพียงในการแก้ไขปัญหาความยากจนอย่างยั่งยืน

3. วิธีการดำเนินการวิจัย

3.1 ประชากรและกลุ่มตัวอย่าง

ประชากรและกลุ่มตัวอย่าง เป็นผู้ที่มิมีภูมิลำเนาอยู่ในเขตพื้นที่ หมู่ที่ 2 บ้านท่ามะขาม ต.ท่ามะขาม อ. เมือง จ. กาญจนบุรี ใช้วิธีเลือกกลุ่มตัวอย่าง คือ วิธีสุ่มไม่เฉพาะเจาะจง เพื่อเลือกกลุ่มตัวอย่างสำหรับให้ข้อมูลทางด้านสิ่งที่ต้องการให้มีในเว็บไซต์ฐานข้อมูลชุมชน

3.2 เครื่องมือที่ใช้ในการวิจัย

1. เว็บไซต์ฐานข้อมูลชุมชน จัดทำขึ้นเพื่อนำเสนอข้อมูลเกี่ยวกับชุมชนให้คนภายนอกและรับรู้ ตลอดจนใช้เป็นแหล่งแลกเปลี่ยนเรียนรู้ของคนในชุมชน ที่สร้างขึ้นโดยใช้ทฤษฎีการออกแบบเน้นความง่ายต่อการเข้าถึงและใช้งานโดยกลุ่มเป้าหมาย โดยในระยะแรกเน้นการพัฒนาและให้ข้อมูลทางด้านการเกษตรเพื่อการพัฒนาด้านเกษตรแบบพอเพียง

2. แบบสอบถามความพึงพอใจการถ่ายทอดความรู้ และการใช้งานเว็บไซต์ฐานข้อมูลชุมชน

3.3 การวิเคราะห์ข้อมูล

วิเคราะห์ข้อมูลจากการดำเนินการวิจัย โดยใช้สถิติพื้นฐาน ได้แก่ ร้อยละ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน โดยนำค่าเฉลี่ยที่ได้ เทียบกับเกณฑ์ประเมิน 5 ระดับ [11] แบ่งการวิเคราะห์เป็น 2 ส่วน คือ

1. วิเคราะห์ข้อมูลที่ได้จากการตอบแบบประเมินการทดลองใช้งานเว็บไซต์ฐานข้อมูล www.thamakham.com เพื่อความคิดเห็นหรือข้อเสนอแนะจากผู้เข้าร่วมทดลองใช้งาน

2. วิเคราะห์ข้อมูลที่ได้จากแบบสอบถามความพึงพอใจในการเข้ารับการถ่ายทอดความรู้เกี่ยวกับฐานข้อมูลออนไลน์ของสมาชิกในชุมชนบ้านท่ามะขาม หมู่ที่ 2 และชุมชนใกล้เคียง โดยทำการถ่ายทอดความรู้ทั้งสิ้น 6 ครั้ง

4. ผลการดำเนินงาน

งานวิจัยนี้แบ่งการดำเนินงานวิจัยเป็น 2 ขั้นตอน คือ จัดทำเว็บไซต์ฐานข้อมูลชุมชน และถ่ายทอดความรู้การใช้งานเว็บไซต์ฐานข้อมูลชุมชน ซึ่งผลที่ได้อยู่ในรูปของ เว็บไซต์ฐานข้อมูลชุมชน และผลวิเคราะห์ความคิดเห็นและความพึงพอใจการใช้งานเว็บไซต์ฐานข้อมูลชุมชนและการถ่ายทอดความรู้การใช้งานเว็บไซต์เพื่อการพัฒนาเศรษฐกิจภาคเกษตรแบบพอเพียงที่ยั่งยืน

4.1 เว็บไซต์ฐานข้อมูลชุมชน

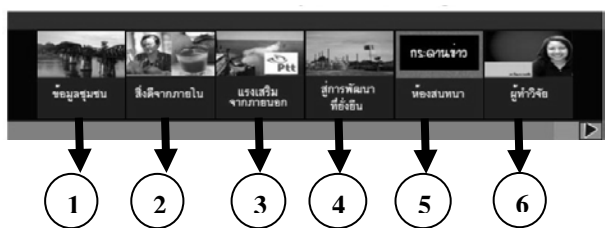
เว็บไซต์ฐานข้อมูลชุมชนที่จัดทำขึ้น มีวัตถุประสงค์เพื่อนำเสนอข้อมูลเกี่ยวกับชุมชนให้คนภายนอกและรับรู้ ตลอดจนใช้เป็นแหล่งแลกเปลี่ยนเรียนรู้ของคนในชุมชน เพื่อการพัฒนาทางด้านการเกษตรแบบพอเพียงที่ยั่งยืน ซึ่งเน้นการออกแบบเว็บไซต์ที่ใช้งานง่าย ลดความซับซ้อนของระดับการเข้าถึงข้อมูลด้วยการนำเสนอข้อมูลทั้งหมดโดยภาพรวมผ่านหมวดหมู่ทั้ง 6 ที่จัดแบ่งไว้ และจัดทำลิ้งค์ข้อมูลแบบไฟล์แนบแบบ .PDF เพื่อให้เปิดอ่านหรือดาวน์โหลดไปศึกษาต่อได้ ผู้สนใจสามารถเข้าชมได้ผ่านเว็บไซต์ www.thamakham.com หน้าแรกของเว็บไซต์ แสดงดังภาพที่ 1



ภาพที่ 1: หน้าแรกของเว็บไซต์ www.thamakham.com

เนื้อหาที่นำเสนอภายในเว็บไซต์แบ่งหมวดหมู่ ออกได้ 6 กลุ่ม คือ 1. ข้อมูลชุมชน 2. สิ่งดีจากภายใน 3. แรงเสริมจากภายนอก 4. ผู้การพัฒนาที่ยั่งยืน 5. กระดานข่าว และ 6. ผู้วิจัย

ซึ่งผู้สนใจสามารถคลิกลิงค์เข้าไปยังรูปภาพ ในลำดับเลข 1-6 ดังแสดงในภาพที่ 2 เพื่อเข้าไปยังหน้าเว็บของแต่ละหมวดหมู่



ภาพที่ 2: ภาพที่ใช้เชื่อมโยงเข้าสู่หน้าเว็บของแต่ละหมวดหมู่

1. ข้อมูลชุมชน นำเสนอข้อมูลเพื่อสะท้อนสิ่งต่างๆ ที่มีลักษณะเป็นรูปธรรม ที่เกี่ยวข้องและบอกความเป็นตัวตนของบ้านท่ามะขาม หมู่ที่ 2 ต.ท่ามะขาม อ.เมือง จ. กาญจนบุรี ประกอบด้วยข้อมูลเกี่ยวกับ ความเป็นมา ขนาด ที่ตั้ง ประชากร ผู้นำหมู่บ้าน และองค์กรในชุมชนที่มีลักษณะเป็นสิ่งปลูกสร้าง
2. สิ่งดีจากภายใน นำเสนอข้อมูลเพื่อสะท้อนศักยภาพของชุมชนในด้านต่างๆ ให้กับคนภายนอกได้รับรู้ ประกอบด้วย ศักยภาพด้านการศึกษา การประกอบอาชีพ ผู้มีความรู้/ภูมิปัญญาท้องถิ่น การรวมกลุ่มของสมาชิกในชุมชนเพื่อดำเนินการกิจกรรมต่างๆ ความต้องการ ของคนในชุมชน การเป็นคนต้นแบบและบุคคลตัวอย่างที่ได้รับการยอมรับจากภายนอก
3. แรงเสริมจากภายนอก นำเสนอข้อมูลเพื่อให้เห็นว่าบ้านท่ามะขาม หมู่ที่ 2 ได้รับการสนับสนุนเพื่อการพัฒนาในด้านต่างๆ ในรูปแบบใดบ้าง ทั้งในรูปของงบประมาณ เงินทุนสนับสนุน งานวิจัย กิจกรรม จากหน่วยงานต่างๆ ทั้งภาครัฐและเอกชน
4. สู่การพัฒนาที่ยั่งยืน นำเสนอข้อมูลเกี่ยวกับแนวทางการปฏิบัติที่จะนำชุมชนไปสู่การพัฒนาที่ยั่งยืน โดยเน้นการนำหลักปรัชญาเศรษฐกิจพอเพียงควบคู่กับการใช้เทคโนโลยีสืบค้นข้อมูลและอินเทอร์เน็ตผ่านเว็บไซต์ที่มีอยู่มากมายในปัจจุบันมาใช้เพื่อการพัฒนาด้านการเกษตรแบบยั่งยืนเป็นเบื้องต้น
5. กระดานข่าว เป็นพื้นที่ซึ่งให้ทั้งสมาชิกในชุมชนและผู้สนใจ ได้เข้ามาแลกเปลี่ยนข้อมูลข่าวสาร โดยบุคคลทั่วไปสามารถเข้ามาอ่านข้อมูลได้ แต่ถ้าต้องการจะเขียน

ข้อความโต้ตอบ หรือตั้งกระทู้ใหม่ ต้องสมัครเป็นสมาชิกเสียก่อน ทั้งนี้ในการสมัครสมาชิกไม่ต้องเสียค่าใช้จ่ายใดๆ

6. ผู้ทำวิจัย นำเสนอเกี่ยวกับผู้วิจัย ทั้งประวัติงานวิจัย ประวัติการทำงาน สิ่งที่น่าสนใจ รวมถึงข้อสรุปถึงสิ่งที่ได้จากการทำงานวิจัยชิ้นนี้

4.2 ผลวิเคราะห์ความคิดเห็นและความพึงพอใจ

ผลวิเคราะห์ความคิดเห็นและความพึงพอใจ แบ่งได้เป็น 2 ส่วน คือ 1. ผลประเมินจากการใช้งานเว็บไซต์ฐานข้อมูลชุมชน และ 2. ผลความพึงพอใจจากผู้รับการถ่ายทอดความรู้การใช้งานเว็บไซต์เพื่อการพัฒนาเศรษฐกิจภาคเกษตรแบบพอเพียงที่ยั่งยืนรายละเอียดมีดังนี้

1. ผลประเมินจากการใช้งานเว็บไซต์ฐานข้อมูลชุมชน www.thamakhom.com ภายใต้โครงการ การจัดการระบบฐานข้อมูลชุมชนแบบออนไลน์ เพื่อการพัฒนาเศรษฐกิจภาคเกษตรแบบพอเพียงที่ยั่งยืน กรณีศึกษาชุมชนบ้านท่ามะขาม รายละเอียดมีดังนี้

มีเข้าร่วมทดลองและประเมินผลจำนวนทั้งสิ้น 92 คน เป็นหญิง 45 คน เป็นชาย 47 คน มีอายุตั้งแต่ต่ำกว่า 18 ปี ถึงอายุ 55 ปี มีการศึกษาระดับปริญญาตรีถึงปริญญาโท ประกอบอาชีพนักเรียน/นักศึกษา ข้าราชการ พนักงานของรัฐ/รัฐวิสาหกิจ และเป็นลูกจ้าง โดยทั้งหมดมีความเห็นว่า คอมพิวเตอร์มีความจำเป็นหรือมีผลต่ออาชีพการทำงาน และสามารถใช้งานระบบอินเทอร์เน็ตเพื่อการสืบค้นได้ ในประเด็นการเข้าเยี่ยมชมหรือสืบค้นข้อมูลเกี่ยวกับบ้านท่ามะขามจังหวัดกาญจนบุรีจากเว็บไซต์ ผู้ประเมินสองในสามตอบว่าเคยเข้าใช้ ส่วนที่เหลือตอบว่าไม่เคย

ผู้ตอบแบบประเมินมีลักษณะการใช้งานระบบเครือข่ายอินเทอร์เน็ตในการทำกิจกรรมต่างๆ เฉลี่ยสัปดาห์ละ 4 ครั้ง เพื่อทำกิจกรรมต่างๆ และมีความคิดเห็นในประเด็นหลักๆ แสดงในตารางที่ 1 นอกจากนี้ ยังมีข้อเสนอแนะอื่นๆ จากผู้ทดลองใช้และประเมินการใช้งานเว็บไซต์ ในเรื่องของการจัดรูปแบบ การจัดการภาพ การเพิ่มช่องทางการติดต่อ หรือเพิ่มเติมรูปแบบสื่อที่นำเสนอ เพื่อให้เว็บไซต์ทำงานได้อย่างมีประสิทธิภาพและตรงตามวัตถุประสงค์งานวิจัยมากที่สุด

ตารางที่ 1: ความคิดเห็นของผู้ใช้ต่อการใช้งานเว็บไซต์

ประเด็นข้อคิดเห็นหลัก	ค่าเฉลี่ย
ด้านข้อมูลและเนื้อหา	4.12
ด้านการออกแบบและการจัดรูปแบบของเว็บไซต์	4.05
ด้านประโยชน์และการนำไปใช้	4.20

2. ผลการประเมินความพึงพอใจจากผู้เข้ารับการถ่ายทอดความรู้การใช้งานเว็บไซต์เพื่อการพัฒนาเศรษฐกิจภาคเกษตรแบบพอเพียงที่ยั่งยืนจำนวนการถ่ายทอดทั้งสิ้น 6 ครั้ง มีผู้เข้ารับการถ่ายทอดทั้งสิ้น 150 คน มีผู้ตอบแบบสอบถาม 145 คน ผู้เข้ารับการถ่ายทอดมีทั้งอยู่ในชุมชนท่ามะขาม หมู่ที่ 2 และชุมชนใกล้เคียง รายละเอียดวัน เวลา สถานที่ สำหรับการถ่ายทอดแสดงไว้ดังตารางที่ 2 และผลที่ได้จากการถ่ายทอดในรูปแบบของความพึงพอใจในประเด็นต่างๆ ของการถ่ายทอดทั้ง 6 ครั้ง แสดงดังตารางที่ 3

ตารางที่ 2: รายละเอียด วัน เวลา สถานที่ และผู้ร่วมรับการถ่ายทอด

ครั้งที่	สถานที่/เวลา	ผู้เข้าร่วม
1	มหาวิทยาลัยราชภัฏกาญจนบุรี 2 กันยายน 2555 เวลา 9.00 – 12.30 น.	31 คน
2	มหาวิทยาลัยราชภัฏกาญจนบุรี 2 กันยายน 2555 เวลา 13.00 – 16.30 น.	18 คน
3	บ้านเลขที่ 16/1 หมู่ที่ 2 ต.ท่ามะขาม อ.เมือง จ.กาญจนบุรี 8 กันยายน 2555 เวลา 17.30 – 21.00 น.	29 คน
4	โรงเรียนวัดท่ามะขาม 10 กันยายน 2555 เวลา 9.00 – 12.30 น.	21 คน
5	โรงเรียนวัดท่ามะขาม 10 กันยายน 2555 เวลา 13.00 – 16.30 น.	29 คน
6	ศูนย์กิจกรรมธรรมชาติท่ามะขาม เลขที่ 213 หมู่ที่ 2 ต.ท่ามะขาม อ.เมือง จ.กาญจนบุรี 14 กันยายน 2555 เวลา 13.00 – 16.30 น.	17 คน

ตารางที่ 3: ความพึงพอใจของผู้เข้ารับการถ่ายทอด 6 ครั้ง

ประเด็นความพึงพอใจ	ค่าเฉลี่ยความพึงพอใจการถ่ายทอดแต่ละครั้ง					
	1	2	3	4	5	6
วิทยากร	4.06	4.06	4.23	4.29	4.43	4.04
สถานที่/ระยะเวลา	4.06	4.37	4.16	4.22	4.36	4.08
ความรู้ความเข้าใจ	4.31	4.31	4.04	4.33	4.14	3.91
การนำความรู้ไปใช้	3.77	4.15	4.05	4.55	4.41	4.07

ซึ่งจากการถ่ายทอดแต่ละครั้ง ได้มีผู้ให้ข้อเสนอแนะในประเด็นที่หลากหลาย สรุปเป็นข้อๆ ได้ดังนี้

1. อยากให้ดำเนินการต่อไปเรื่อยๆ และสนับสนุนให้ทุกคนได้สนใจในเรื่องระบบการสื่อสารที่ทันต่อโลก
 2. ผู้เข้ารับการถ่ายทอดมีพื้นฐานไม่เท่ากัน อยากให้เริ่มต้นปูพื้นฐานตั้งแต่การใช้งานคอมพิวเตอร์เบื้องต้น
 3. หน้าเว็บไซต์น่าสนใจ สามารถเข้าถึงเนื้อหาได้ง่าย
 4. กลุ่มผู้อาวุโสเห็นว่าตัวอักษรบนจอเล็กไป
 5. ต้องการข้อมูลเพิ่มเติมด้านหนี้สิน และต้องการได้รับการถ่ายทอดความรู้เรื่องการใช้งานคอมพิวเตอร์ สืบค้นอินเทอร์เน็ต
 6. ยินดีที่มีเว็บไซต์สำหรับชุมชนเพื่อเผยแพร่ข้อมูลและแลกเปลี่ยนเรียนรู้
 7. อยากให้มาสอน มาอบรม และมาร่วมถ่ายทอดอีก เพราะได้ความรู้เกี่ยวกับคอมพิวเตอร์ และอยากให้เพิ่มเวลาขึ้น
- สิ่งที่ทำให้ผู้วิจัยรู้สึกทึ่งว่างานชิ้นนี้เดินมาถูกทางแล้ว คือ การที่ได้รับทราบจากผู้เข้ารับการถ่ายทอดในครั้งที่ 3 เมื่อวันที่ 8 กันยายน 2555 รวมถึงชาวชุมชนท่ามะขาม หมู่ที่ 2 กลุ่มอื่นๆ ได้พูดคุยกันในวงกว้าง ถึงข้อมูลต่างๆ ที่จะนำไปแสดงบนเว็บไซต์ www.thamakham.com เพื่อให้ผู้รับข่าวสารได้รับข้อมูล ข้อเท็จจริง ที่เป็นปัจจุบันที่สุด ทำให้เกิดการทบทวนกันเองภายในหมู่บ้าน ถึงสิ่งที่มี สิ่งที่เป็นอยู่ รวมไปถึง สิ่งที่กำลังดำเนินการและต้องดำเนินการของชุมชน

5. สรุป

งานวิจัยนี้ ได้รวบรวมข้อมูลของชุมชนบ้านท่ามะขามหมู่ที่ 2 ไว้ในเว็บไซต์ www.thamakham.com เพื่อให้บุคคลทั้งภายในและภายนอกชุมชนได้รับรู้ข้อมูล ทั้งข้อมูลทั่วไปว่าตั้งอยู่ที่ไหน ประชากรเท่าไร มีอะไรเด่น ซึ่งอาจเป็นได้ทั้ง

สถานที่ และความรู้ความสามารถของคนในชุมชน นอกจากนี้ ข้อมูลของงบประมาณสนับสนุนทั้งจากภาครัฐและเอกชนยังช่วยให้คนในชุมชนมองเห็นแนวทางในการพัฒนาทั้งด้านพื้นที่ และอาชีพจากงบประมาณที่มีอยู่ หรือมองเห็นโอกาสในการสร้างความร่วมมือกับแหล่งทุนภายนอก ซึ่งได้รวบรวมบางส่วนไว้ในเว็บไซต์ ผู้สนใจจากภายนอกเองก็สามารถศึกษาข้อมูลผ่านเว็บไซต์ของชุมชน โดยติดต่อผู้ที่มีส่วนเกี่ยวข้องหรือรับผิดชอบได้โดยตรง ซึ่งช่วยเพิ่มโอกาสให้กับชุมชน และเพื่อให้เกิดความยั่งยืนเป็นประโยชน์อย่างแท้จริง การปรับปรุงข้อมูลเว็บไซต์ต้องกระทำอย่างต่อเนื่องและเกิดขึ้น โดยความร่วมมือของคนในชุมชนและผู้ที่เกี่ยวข้อง

จากการออกแบบและทดลองใช้งานระบบ ผู้ประเมินการใช้งาน ที่ส่วนใหญ่มีความรู้ทางด้านคอมพิวเตอร์และอินเทอร์เน็ต โดยมีความคิดเห็น ว่าข้อมูลและเนื้อหา การออกแบบและการจัดรูปแบบของเว็บไซต์ และประโยชน์และการนำไปใช้ ว่ามีคุณภาพในระดับมาก ซึ่งข้อเสนอแนะต่างๆ ในเรื่องของการจัดรูปแบบ การจัดการภาพ การเพิ่มช่องทางการติดต่อ หรือเพิ่มเติมรูปแบบสื่อที่นำเสนอ ได้ถูกนำมาปรับใช้กับหลักการออกแบบเว็บไซต์สำหรับกลุ่มเป้าหมาย คือ ชาวบ้าน ส่วนส่วนใหญ่มีทักษะทางด้านคอมพิวเตอร์ค่อนข้างน้อย เพื่อให้เว็บไซต์ทำงานได้อย่างมีประสิทธิภาพและตรงตามวัตถุประสงค์งานวิจัยมากที่สุด

ผลจากการถ่ายทอดความรู้เกี่ยวกับเว็บไซต์ชุมชน www.thamakham.com ทั้ง 6 ครั้งนั้น ผู้เข้ารับการถ่ายทอดพึงพอใจในระดับมาก ให้ข้อเสนอแนะและข้อมูลต่างๆ ที่เป็นประโยชน์ในการปรับปรุงเว็บไซต์ ข้อมูล และวิธีการนำเสนอข้อมูล รวมไปถึงขยายผลการถ่ายทอดและพูดคุยเรื่องราวของชุมชนท่ามะขามหมู่ 2 ที่นำเสนอในรูปแบบเว็บไซต์กันอย่างกว้างขวาง และนำข้อมูลรายละเอียดที่นำเสนอในเว็บไซต์มาพูดคุยกับนอกรอบเพื่อให้เกิดข้อสรุปว่าต้องการให้นำเสนอสิ่งใดเพิ่มเติม และสิ่งใดที่ต้องปรับแก้ นอกจากนี้ยังพบว่ามีกิจกรรมต่างๆ ภายในชุมชนที่สามารถต่อยอดงานวิจัยได้อีก ยกตัวอย่างเช่น พลังงานชุมชนทางเลือก เกษตรอินทรีย์ รวมถึงการจัดทำโครงการถ่ายทอดความรู้ทางด้านคอมพิวเตอร์และภาษาอังกฤษเบื้องต้น เป็นต้น ที่สำคัญการพูดคุยกับชาวบ้านในชุมชน ทำให้ทราบถึงปัญหา ความต้องการพัฒนาตัวเองเพื่อให้

เกิดความยั่งยืนทางความรู้ และประสานความร่วมมือกับหน่วยงานการศึกษามาใช้เพื่อสร้างความร่วมมือให้เกิดการพัฒนาที่ยั่งยืน เช่น การให้ความรู้ในเรื่องภาษาอังกฤษ คอมพิวเตอร์ รวมถึงการเตรียมความพร้อมเพื่อพัฒนาอาชีพในด้านที่สนใจ

เอกสารอ้างอิง

- [1] กรมการพัฒนาชุมชน, รายงานหมู่บ้านชนบทไทย จากข้อมูลพื้นฐานระดับหมู่บ้าน (กชช 2 ก.) ปี 2555, กรมการพัฒนาชุมชน กระทรวงมหาดไทย, 2555.
- [2] นิรุบล ทองชัย, การจัดการระบบฐานข้อมูลชุมชนแบบออนไลน์ เพื่อการพัฒนาเศรษฐกิจภาคเกษตรแบบพอเพียงที่ยั่งยืน กรณีศึกษาชุมชนบ้านท่ามะขาม จังหวัดกาญจนบุรี, สำนักงบประมาณ, 2555.
- [3] ครู กศน. การเข้าถึงชุมชนและการจัดทำฐานข้อมูลชุมชน ค้นเมื่อ มกราคม, 12, 2555 จาก <http://www.gotoknow.org/blogs/posts/403949>
- [4] สุทธิศา จิระนันท์ และ ชีรวัดน์ หังสพฤกษ์. “ระบบฐานข้อมูลบนเว็บเพื่อนำเสนอข่าวชายแดนภาคใต้” วารสารเทคโนโลยีสารสนเทศ ปีที่ 3 ฉบับที่ 5 2550 หน้า 34-42.
- [5] วรณพร ทีเก่ง, รายงานวิจัยระบบฐานข้อมูลแพทย์พื้นบ้าน จ.ตาก, มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา จ.ตาก, 2551.
- [6] ผุสดี จิระวัฒนกิจ, เอกสารประกอบการจัดทำฐานข้อมูลท้องถิ่น. สพท. นครปฐม เขต 1, 2553.
- [7] อัญชนา บุญเรือง, การศึกษาเปรียบเทียบพฤติกรรมการใช้เทคโนโลยีสารสนเทศของหญิงและชายในองค์กรธุรกิจ, จุฬาลงกรณ์: กรุงเทพฯ มหาวิทยาลัย, 2540.
- [8] มูลนิธิชัยพัฒนา, เศรษฐกิจพอเพียงและทฤษฎีใหม่, มูลนิธิชัยพัฒนา: กรุงเทพฯ, 2550.
- [9] ทศนีย์ เศรษฐ์บุญสร้าง, เอกสารประกอบการอบรมสัมมนา เรื่อง เกษตรอินทรีย์องค์รวม (Holistic Organic Farming). ค้นเมื่อ มกราคม, 7, 2555, จาก <http://www.dnaf.org/news/data-006.html>
- [10] ไชยรัตน์ ปรานี, การวิจัยเอกสาร ชุมชนต้นแบบที่นำแนวทางพระราชดำริเรื่องเศรษฐกิจพอเพียง ไปใช้ในการแก้ปัญหาความยากจนอย่างยั่งยืน, สถาบันวิจัยและพัฒนา มหาวิทยาลัยราชภัฏนครสวรรค์ จังหวัดนครสวรรค์, 2551.
- [11] บุญชม ศรีสะอาด, การวิจัยเบื้องต้น, สุวีริยาสาส์น: กรุงเทพฯ , 2535.

ระบบตรรกศาสตร์คลุมเครือชนิดที่ 2 สอนด้วยขั้นตอนวิธีแบบผสมผสานเชิงสามัญสำนึก สำหรับตลาดหลักทรัพย์แห่งประเทศไทย

Interval Type-2 Fuzzy Systems Trained by Hybrid Heuristic Algorithm for Stock Exchange of Thailand

อดิศักดิ์ แสงส่องฟ้า (Adisak Sangsongfa) และ พยุง มีสัจ (Phayung Meesad)

คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

Adisak.sangsongfa@gmail.com, pym@kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้นำเสนอวิธีการสอนให้ ระบบตรรกศาสตร์แบบคลุมเครือชนิดที่ 2 ทำการพยากรณ์อนุกรมเวลาของราคาตลาดหลักทรัพย์แห่งประเทศไทย ด้วยขั้นตอนวิธีแบบผสมผสานเชิงสามัญสำนึกแบบใหม่ นั่นคือ ITFS2HH ซึ่งเป็นการรวมกันระหว่างขั้นตอนวิธีเชิงพันธุกรรม และการเคลื่อนที่ของกลุ่มอนุภาค โดยใช้ข้อมูลตั้งแต่ปี ค.ศ. 2010-2013 เป็นจำนวน 756 วัน แบ่งเป็น 456 วัน สำหรับชุดสอน และ 300 วัน สำหรับชุดทดสอบ ผลการสร้างระบบฟัซซีมีกฎ 10 กฎ มีค่าผิดพลาดสำหรับชุดสอน $RMSE = 0.03982$ และมีค่าผิดพลาด $RMSE = 0.05174$ สำหรับชุดทดสอบ

คำสำคัญ: ตรรกศาสตร์คลุมเครือชนิดที่ 2, วิธีแบบผสมผสานเชิงสามัญสำนึก, การพยากรณ์อนุกรมเวลา

Abstract

In this research, an interval type-2 fuzzy system trained by a new hybrid heuristic algorithm using genetic algorithm and particle swarm optimization (ITFS2HH) is proposed for time series prediction. The Stock Exchange of Thailand (SET) data set between the years 2010-2013 was used. The author selected 756 days: 456 days for training and 300 for testing. A fuzzy system with 10 rules was generated and had $RMSE = 0.03982$ in the training set. The resulted fuzzy system produced $RMSE = 0.05174$ on the test set.

Keyword: Interval Type-2 Fuzzy Logic System, New Hybrid Heuristic Algorithm, Time Series Prediction

1. บทนำ

การพยากรณ์อนุกรมเวลาของตลาดหลักทรัพย์นับเป็นเรื่องที่น่าสนใจและท้าทายต่อนักลงทุนและนักวิชาการ เพราะตลาดหลักทรัพย์มีความผันผวนและอ่อนไหวต่อหลายๆปัจจัย เช่น ปัจจัยความน่าเชื่อถือต่อการลงทุน ทางด้านการขยายตัวต่อการลงทุน ทางด้านสภาพแวดล้อมทางเศรษฐกิจ และอื่นๆอีกมากมาย โดยเฉพาะตลาดหลักทรัพย์แห่งประเทศไทย (Stock Exchange of Thailand (SET)) จะมีปัจจัยที่ต้องพิจารณาเพิ่มเติมที่นอกเหนือไปจากปัจจัยต่างๆตามที่ได้กล่าวมาแล้วในข้างต้น เช่น ปัจจัยจากตลาดหุ้น Dow Jones Index (New York), Strats Times Index (Singapore), Nekkei Index (Japan), Hang Seng Index (Hong Kong), อัตราดอกเบี้ยต่ำ, ราคาทองคำ และอัตราค่าเงินบาท เป็นต้น [1-2]

ในการพยากรณ์อนุกรมเวลาของตลาดหลักทรัพย์แห่งประเทศไทยมีการศึกษาและทำการวิจัยในหลายๆวิธีเพื่อให้ได้ค่าการพยากรณ์ที่มีค่าของ รากของค่าผิดพลาดกำลังสองเฉลี่ย (Root Mean Square Error: RMSE) หรือค่าผิดพลาดกำลังสองเฉลี่ย (Mean Square Error (MSE)) น้อยที่สุดเช่น การพยากรณ์ด้วย Multiple Regression [1], Markov Regime Switching GRAPH [2], Evolution Strategies [3-4], Black-Box[5], An ANFIS Model [6], Neural Network [7] และ Fuzzy Logic System [8] เป็นต้น ซึ่งค่าของความผิดพลาดกำลังสองเฉลี่ย หรือ รากของค่าผิดพลาดกำลังสองเฉลี่ยที่มีค่าน้อยนั้นหมายถึงค่าที่พยากรณ์ได้ก็จะตรงกับค่าที่เป็นตามความเป็นจริง ยิ่งค่าน้อยมากเท่าไรก็จะตรงกับค่าที่เป็นตามความเป็นจริงมากขึ้นเท่านั้น ซึ่งในแต่ละงานวิจัยต่างก็มีข้อแตกต่างกันออกไปสามารถแสดงได้ดังต่อไปนี้

Sopipan [1] ใช้การวิเคราะห์การถดถอยพหุคูณ (Multiple Regression) เพื่อพยากรณ์ค่าที่เป็นตามความเป็นจริงของ 50

บริษัทชั้นนำในตลาดหลักทรัพย์ของประเทศไทย ทำการวิเคราะห์ห้วงองค์ประกอบของตลาดหลักทรัพย์ของไทยด้วยการกำหนดค่าของตัวแปรที่เกิดสหสัมพันธ์กันเองในระดับสูงคือค่าจากตลาดหลักทรัพย์อื่นๆทั่วโลกเช่น Dow Jones Index (New York), Straits Times Index (Singapore), Nikkei Index (Japan), Hang Seng Index (Hong Kong) เป็นต้น ได้ค่าของระดับคะแนน Principal Component Scores เป็น 99.4% ซึ่งเป็นการนำเอาวิธีทางสถิติและการพิจารณาองค์ประกอบอื่นๆ มาช่วยในการพยากรณ์ Sattayatham [2] ใช้การสร้างแบบจำลองของ autoregressive เพื่อหาค่าเฉลี่ยของดัชนีตลาดหลักทรัพย์แห่งประเทศไทย ในแต่ละวันของสัปดาห์ ต่อจากนั้นทำการพยากรณ์ค่าความผันผวนของดัชนีตลาดหลักทรัพย์ ด้วยแบบจำลองกราฟและ Markov Regime Switching ซึ่งผลการพยากรณ์สามารถบอกได้ว่า ในวันศุกร์ของทุกสัปดาห์จะมีความผันผวนมากกว่าวันอื่นๆ โดยยืนยันการพยากรณ์ด้วยค่าทางสถิติต่างๆ Sutheebanjard [3] ใช้ศาสตร์ของปัญญาประดิษฐ์ในการคำนวณเชิงวิวัฒนาการ เรียกว่า ยุทธศาสตร์เชิงวิวัฒนาการ (Evolution Strategies) เพื่อทำการพยากรณ์ตลาดหุ้นไทย ผลลัพธ์ที่ได้คือค่าของค่าเฉลี่ยเปอร์เซ็นต์ความผิดพลาดสัมบูรณ์ (MAPE) มีค่าน้อยกว่า 1 และผู้วิจัยได้นำผลที่ได้ไปเปรียบเทียบกับงานเดิมซึ่งงานเดิมค่าของ ค่าเฉลี่ยเปอร์เซ็นต์ความผิดพลาดสัมบูรณ์ (MAPE) มีค่าน้อยกว่า 2 [4] Akbar [6] ใช้วิธีการพยากรณ์ด้วย ANFIS Model ด้วยการตั้งเป้าสำหรับการพยากรณ์เพื่อการประมาณการการลงทุนในตลาดหุ้น การลดอัตราความเสี่ยงและเพิ่มความถูกต้องในการลงทุน ในการพัฒนา ANFIS Model มีการนำเอา Fuzzy System เข้ามาช่วยในการจัดกลุ่มข้อมูลของตลาดหลักทรัพย์ เพื่อแสดงให้เห็นถึงประสิทธิภาพของอัลกอริทึม จึงทำการเปรียบเทียบกับวิธีการ Artificial Neural Network, Fuzzy C- Mean และ ANFIS ที่ไม่มีการจัดกลุ่ม ผลที่ได้คือค่า RMSE จาก ANFIS ที่พัฒนาขึ้นมีค่าน้อยกว่าวิธีการอื่นๆ Phaisarn [7] ใช้วิธีการโครงข่ายประสาทเทียมแบบการส่งค่าย้อนกลับ (back propagation) เพื่อทำการพยากรณ์ดัชนีของตลาดหลักทรัพย์แห่งประเทศไทย โดยวัดค่าของ MSE และค่าความผิดพลาดของดัชนีตลาดหุ้นมีค่าน้อยกว่า 2% Hector [8] ใช้การทำงานร่วมกันของระบบตรรกศาสตร์คลุมเครือชนิดที่ 1 และ 2 เช่น การใช้ระบบตรรกศาสตร์คลุมเครือชนิดที่ 1 ในการ

กำหนดค่าของ linguistic parameter แล้วทำการแปลง linguistic parameter นั้นไปเป็นระบบตรรกศาสตร์แบบคลุมเครือชนิดที่ 2 ซึ่งผลลัพธ์ที่ออกมาให้เห็นว่าระบบตรรกศาสตร์แบบคลุมเครือชนิดที่ 2 มีค่าทางสถิติที่ดีกว่าอย่างไรก็ตาม จากงานวิจัยที่ได้กล่าวมาแล้วในข้างต้นนั้น ต่างก็ใช้วิธีการที่มีการใช้งานอย่างแพร่หลาย อาจจะมีการปรับปรุงไม่มากนัก แต่ค่าที่ได้จากการพยากรณ์ต่างก็ยังให้ค่า RMSE ที่มากอยู่เหมือนเดิม

สำหรับในงานวิจัยนี้ผู้วิจัยนำเสนอวิธีการพัฒนาขั้นตอนวิธีแบบผสมผสานเชิงสามัญลักษณ์ด้วยขั้นตอนวิธีเชิงพันธุกรรมและการเคลื่อนที่ของกลุ่มอนุภาคเพื่อสอนให้ ตรรกศาสตร์แบบคลุมเครือชนิดที่ 2 (ITFS2HH) ทำการพยากรณ์อนุกรมเวลาของตลาดหลักทรัพย์แห่งประเทศไทย โดยมุ่งให้ผลพยากรณ์ที่ใกล้เคียงกับค่าที่เป็นตามความเป็นจริงให้มากที่สุด วัดด้วยค่า RMSE ที่น้อย ภายได้ข้อจำกัดของการมีจำนวนฟังก์ชันการเป็นสมาชิกน้อย และค่าพารามิเตอร์ที่เหมาะสม

2. วรรณกรรมที่เกี่ยวข้อง

2.1 ขั้นตอนวิธีเชิงพันธุกรรม (Genetic Algorithm: GA)

ขั้นตอนวิธีเชิงพันธุกรรมเป็นการค้นหาคำตอบโดยประมาณของปัญหา ซึ่งอาศัยหลักของทฤษฎีวิวัฒนาการจากชีววิทยา และการคัดเลือกตามธรรมชาติ เพื่อแก้ปัญหาหาค่าที่เหมาะสมที่สุด โดยการแทนคำตอบที่มีอยู่ในลักษณะของโครโมโซม (chromosome) สรุบบนขั้นตอนของ GA จะประกอบไปด้วย 4 ส่วนคือ 1) ประชากรของกลุ่มคำตอบแต่ละชุด ซึ่งแต่ละชุดจะแสดงค่าที่เป็นไปได้ของคำตอบที่ต้องการ 2) ค่าที่เหมาะสม (fitness function) ซึ่งเป็นค่าที่ได้จากการประเมินของกลุ่มคำตอบแต่ละกลุ่มว่าเป็นค่าที่ต้องการหรือเป็นค่าที่ไม่ต้องการ 3) วิธีการคัดเลือก (selection function) คือวิธีการตัดสินใจว่าจะทำการหยิบกลุ่มคำตอบที่ดีจากกลุ่มประชากรในปัจจุบัน เพื่อจะใช้ในการสร้างกลุ่มประชากรใหม่ และ 4) ตัวปฏิบัติการทางพันธุกรรม (genetic operator) เช่น การไขว้เปลี่ยน (crossover) และการกลายพันธุ์ (mutation) เพื่อเป็นการสร้างพื้นที่ใหม่ที่ใช้สำหรับการสร้างกลุ่มประชากรใหม่ [9] ซึ่งวิธีการขั้นตอนวิธีเชิงพันธุกรรม ไม่ใช่เรื่องใหม่เพราะมีการพัฒนามาตั้งแต่ปี 1965 และงานวิจัยต่างๆก็ได้นำไปประยุกต์ใช้

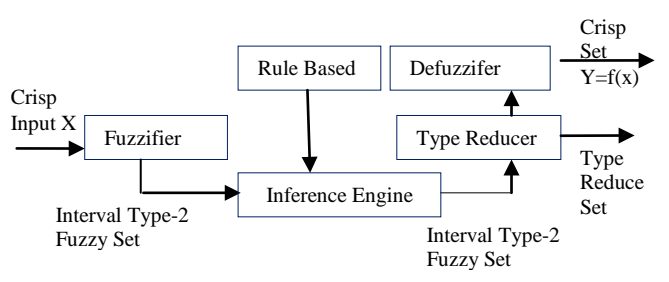
พร้อมทั้งพัฒนาอย่างต่อเนื่อง แต่ก็มีข้อจำกัดของการพัฒนา ทำให้ผลที่ได้ไม่ค่อยเป็นที่น่าพอใจในปัจจุบัน เพราะข้อมูลในปัจจุบันมีขนาดใหญ่และซับซ้อนมากขึ้น [10]

2.2 การเคลื่อนที่ของกลุ่มอนุภาค (Particle Swarm Optimization: PSO)

PSO นับเป็นอัลกอริทึมที่พัฒนาต่อมาจาก GA ซึ่งได้รับการพัฒนาในราวปี 1995 การทำงานของ PSO เลียนแบบจากพฤติกรรมทางสังคมของสิ่งมีชีวิต เช่นการเคลื่อนที่ของฝูงนกที่มีก้ออกหาอาหารพร้อมๆกันเป็นฝูง นกแต่ละตัวจะทำหน้าที่ในการบินค้นหาอาหารและอาจย้ายถิ่นที่อยู่ไปพร้อมๆกัน นกจะสื่อสารกันในกลุ่มเพื่อแลกเปลี่ยนข้อมูลระหว่างกัน ทั้งนี้ นกแต่ละตัวจะเปรียบเสมือนอนุภาคย่อย ที่ประกอบไปด้วยเวกเตอร์สองเวกเตอร์ คือเวกเตอร์ระบุตำแหน่ง (location) และเวกเตอร์ระบุความเร็วในการบิน (velocity) สามารถแสดงด้วยสมการซึ่งแสดงอยู่ในวิธีดำเนินการวิจัย [8]

2.3 ระบบตรรกศาสตร์แบบคลุมเครือชนิดที่ 2

ภาพที่ 1 แสดงแผนภาพการทำงานของระบบตรรกศาสตร์คลุมเครือชนิดที่ 2 ซึ่งแผนภาพจะคล้ายกับระบบตรรกศาสตร์คลุมเครือชนิดที่ 1 แต่จะมีข้อแตกต่างตรงที่แบบที่ 2 มีการนำเอาการประมวลผลค่านำออกมาเพิ่มต่อจากส่วนที่แปลงค่านำออกไปให้อยู่ในช่วงที่เหมาะสม (Defuzzification)



ภาพที่ 1: Interval Type-2 Fuzzy System

จากภาพที่ 1 ในส่วนของ การประมวลผลค่านำออก ในระบบตรรกศาสตร์คลุมเครือชนิดที่ 2 จะประกอบไปด้วย 2 ขั้นตอนคือ ขั้นตอนของ Type reducer block และ Defuzzifier block โดย Type reducer block คือการจับคู่ระหว่าง type -2 fuzzy set ไปเป็น type-1 fuzzy set แล้วทำการ defuzzification ต่อจากนั้นก็ทำการแปลงไปเป็น ค่าของ crisp output [6]

2.4 ตลาดหลักทรัพย์แห่งประเทศไทย (The Stock Exchange of Thailand – SET)

ตลาดหลักทรัพย์แห่งประเทศไทยจัดตั้งขึ้นเมื่อวันที่ 30 เมษายน พ.ศ. 2518 ตามพระราชบัญญัติตลาดหลักทรัพย์แห่งประเทศไทย พ.ศ. 2517 อยู่ภายใต้การกำกับดูแลของสำนักงานคณะกรรมการกำกับหลักทรัพย์และตลาดหลักทรัพย์ (ก.ล.ต.) ทำหน้าที่เป็นตลาดรอง เพื่อแลกเปลี่ยนซื้อขายตราสารทุนของบริษัทต่างๆ ที่ขึ้นทะเบียนไว้และเพื่อให้สามารถระดมเงินทุนเพิ่มเติมจากสาธารณะได้โดยสะดวก ปัจจุบันการดำเนินงานของตลาดหลักทรัพย์ของไทยอยู่ภายใต้พระราชบัญญัติหลักทรัพย์และตลาดหลักทรัพย์ พ.ศ. 2535 [2]

3. วิธีดำเนินการวิจัย

ในการพัฒนาขั้นตอนวิธีแบบผสมผสานเชิงสามัญสำนึกด้วยขั้นตอนวิธีเชิงพันธุกรรม และการเคลื่อนที่ของกลุ่มอนุภาคเพื่อหาค่าพารามิเตอร์ที่เหมาะสมของตรรกศาสตร์แบบคลุมเครือชนิดที่ 2 (ITFS2HH) เพื่อให้ได้ค่านำออกที่เป็นค่าการพยากรณ์อนุกรมเวลาของตลาดหลักทรัพย์แห่งประเทศไทย ที่ใกล้เคียงกับค่าที่เป็นตามความเป็นจริง ให้มากที่สุด ด้วยค่าของ RMSE ที่น้อยที่สุด ดังนั้นผู้วิจัยได้นำเสนอกรอบงานของ ITFS2HH ดังมีรายละเอียด ในภาพที่ 2

จากคำสั่งเทียม (pseudo code) ตามภาพที่ 2 สามารถอธิบายเป็นลำดับขั้นตอนการทำงานโดยแบ่งได้เป็นจำนวน 5 ขั้นตอนดังนี้

- 1) ทำการจัดกลุ่มของข้อมูล เพื่อหาแกนพีชชี โดยใช้วิธีการจัดกลุ่มแบบพีชชี ซีมิน (Fuzzy C-Means) ซึ่งมีสมการการคำนวณความเป็นสมาชิก U ของข้อมูลทุกตัวดังสมการที่ 1

$$U_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{d_{ij}}{d_{kj}} \right)^{\frac{2}{m-1}}} \tag{1}$$

เมื่อ m คือค่าระดับความเป็นสมาชิก, C คือจำนวนกลุ่ม, d_{ij} คือการคำนวณระยะระหว่างกลุ่ม i กับข้อมูล j ตามสูตรของ Euclidean Distance

$$d_{ij} = \|c_i - x_j\| \tag{2}$$

c_i คือ ตัวแทนของกลุ่มที่ i , x_j คือข้อมูลตัวที่ j , สมการ คำนวณค่าดัชนีวัดความผิดพลาด

- Load SET dataset
- Initialize the centering of data and rule base from data clustering using Fuzzy C-mean
- Create a random initial population (Chrom)

[iter=0]

Repeate

- Evaluate each individual in the population using the fitness (Eq. 5)
- Order the population
- Select two chromosome better fitness for crossover 50% and mutation 1% (Eq. 8 and 9)
- Acceptation place new offspring and replace use new generated population

Until(iter ≥ maximum number of generation)

- Create particle from best individual from GA 20% and random particle 80%
- Calculate fitness (Eq.5) for each particle

[iter=0]

Repeat

If (fitness > pBest)

- Assign current fitness as new pBest

Else

- Keep previous pBest

End if

- Assign best particle's pBest to gBest
- Calculate velocity (Eq. 6) used (Eq. 8 and 9.) for inertia and updated location (Eq. 7)
- Select two chromosome better fitness for crossover 50% and mutation 3% (Eq. 8 and 9)
- Acceptation place new offspring and replace use new generated population

Until(iter ≥ maximum number of generation)

- Best individual minimum fitness (Eq.5)

ภาพที่ 2: คำสั่งเทียมแสดงขั้นตอนวิธีแบบผสมผสานเชิงสามัญสำนึก ด้วยขั้นตอนวิธีเชิงพันธุกรรมและการเคลื่อนที่ของกลุ่มอนุภาคเพื่อหาค่าพารามิเตอร์ที่เหมาะสมของตรรกศาสตร์แบบคลุมเครือชนิดที่ 2

min objective function และจัดข้อมูลเข้ากลุ่ม คือ

$$J_m = \sum_{i=1}^c \sum_{j=1}^n U_{ij}^m \|c_i - x_j\|^2, 1 \leq m < \infty \quad (3)$$

และสมการการคำนวณตัวแทนของกลุ่มใหม่ คือ

$$C_i = \frac{\sum_{j=1}^n U_{ij}^m x_j}{\sum_{j=1}^n U_{ij}^m} \quad (4)$$

2) ทำการสุ่มประชากรและกำหนดโครโมโซม (chromosome) ของฝูง (swarm) และ กลุ่ม (particle) เพื่อหา position และ velocity ดังมีโครโมโซมดังนี้

$$Chrom = [IL, IM, R]$$

โดยที่ $IL = [IL_1, IL_2, \dots, IL_n]$ คือเซตของ interval linguistic variable, $IM = [im_{11}, im_{12}, \dots, im_{n,IL_n}]$ คือเซตของ interval membership function parameter ของ interval linguistic variable, $R = [R_1, R_2, \dots, R_{IL_1 \times IL_2, \dots, IL_n}]$ คือ fuzzy rule และ R_1 คือ ตัวเลขที่เป็นจำนวนเต็มที่ได้จากการสร้าง fuzzy rule ของแต่ละมิติใน linguistic variable ดังแสดงใน ตารางที่ 1

ตารางที่ 1: ตารางแสดงการกำหนดค่าเริ่มต้นของ ฟังก์ชันการเป็นสมาชิก สำหรับตัวแปร (Linguistic Variable) 5 ตัวแปร

Linguistic Index	Linguistic Terms
1	Very Low
2	Low
3	Medium
4	High
5	Very High

3) กำหนดสมการในการหาค่าของ Fitness Function โดยใช้ ค่าความคลาดเคลื่อนของรากกำลังสองเฉลี่ย (Root Mean Square Error (RMSE)) ดังมีสมการดังนี้

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (X_{obs,i} - X_{model,i})^2}{n}} \quad (5)$$

โดยที่ X_{obs} คือค่าที่ได้จากการสังเกต X_{model} คือค่าที่ได้จากแม่แบบในเวลาที i

4) กัดเลือกกลุ่มประชากรของฝูง ด้วยการใช้วิธีการของขั้นตอนวิธีเชิงพันธุกรรม ด้วยอัตราการแลกเปลี่ยนยีน (crossover) ที่ 50% และ อัตราการกลายพันธุ์ (mutation) ที่ 1% จะทำให้ได้ประชากรที่เป็นชนชั้นนำ (elite) จากนั้นนำประชากรที่เป็นชนชั้นนำเข้าร่วมกับกลุ่มประชากรทั่วไปในอัตราชนชั้นนำจำนวน 20% ประชากรทั่วไป 80%

5) ปรับปรุงค่าของ position และ velocity ของกลุ่มประชากร พร้อมทั้งปรับปรุงคุณภาพของกลุ่มประชากรด้วยการใช้การกระทำของขั้นตอนวิธีเชิงพันธุกรรม ด้วยอัตราการแลกเปลี่ยนยีน 50% การกลายพันธุ์ที่ 3% ซึ่งสมการมีดังนี้

$$V_{i(t+1)} = wV_{i(t)} + c_1r_1(P_i - x_{i(t)}) + c_2r_2(P_g - x_{i(t)}) \quad (6)$$

$$X_{i(t+1)} = X_{i(t)} + V_{i(t+1)} \quad (7)$$

โดยที่ P_i คือ pbest, P_g คือ gbest, c_1, c_2 คือค่าคงที่ของความเร่ง (acceleration) ที่มีช่วงอยู่ระหว่าง $[0..2]$, r_1, r_2 คือค่าตัวเลขที่ได้จากการสุ่มของตัวเลข $[0, 1]$ และ w คือค่าที่ได้จากการถ่วงน้ำหนักของความเฉื่อย (inertia) ซึ่งได้จากสมการของ Chaotic inertia weight ดังสมการ

$$w = (w_1 - w_2) \times \frac{Max_{iter} - iter}{Max_{iter}} + w_2 \times z \quad (8)$$

$$z = 4 \times z \times (1 - z) \quad (9)$$

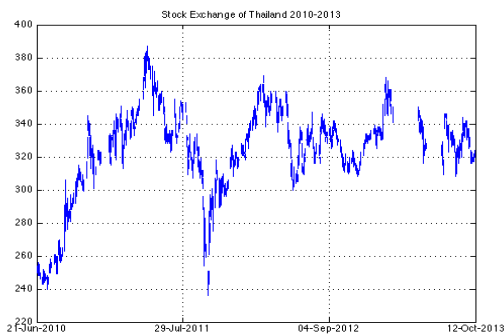
โดยที่ $w_1=0.9$, $w_2=0.4$ และ z คือค่าที่ได้จากการสุ่มระหว่าง $0..1$

4. ผลการทดลอง

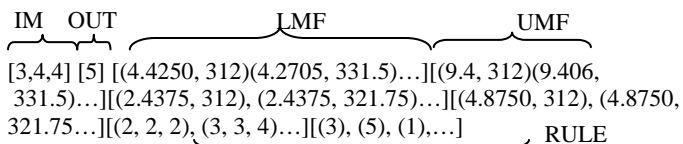
จากการพัฒนา ITFS2HH เพื่อการพยากรณ์อนุกรมเวลาของตลาดหลักทรัพย์แห่งประเทศไทยสามารถอธิบายได้ดังนี้คือ

4.1 ข้อมูล

ผู้วิจัยได้ใช้ข้อมูลของตลาดหลักทรัพย์แห่งประเทศไทย [9] ซึ่งเป็นข้อมูลการซื้อขายหลักทรัพย์ที่เกี่ยวกับน้ำมัน ตั้งแต่ปี ค.ศ. 2010-2013 ดังแสดงในภาพที่ 3 โดยในการวิจัย ผู้วิจัยได้แบ่งข้อมูลสำหรับใช้ในการฝึกสอน จำนวน 456 วัน โดยเริ่มตั้งแต่เดือน กรกฎาคม ปี 2010 จนถึงเดือน กรกฎาคม ปี 2012 เพื่อที่จะใช้สำหรับทดสอบและพยากรณ์ข้อมูลในปี 2012-2013 เป็นจำนวน 300 วัน โดยเริ่มตั้งแต่เดือน เมษายน ปี 2012 จนถึงเดือน พฤศจิกายน ปี 2013



ภาพที่ 3: Stock Exchange of Thailand 2010-2013



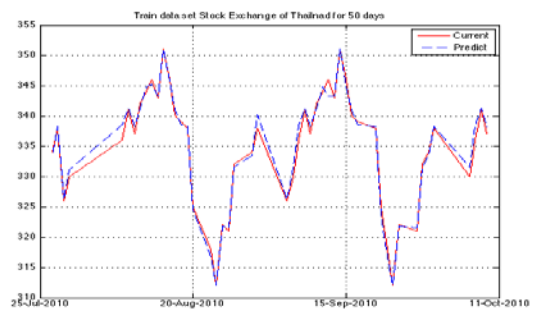
ภาพที่ 4: ITFS2HH System

4.2 ผลการทดลอง

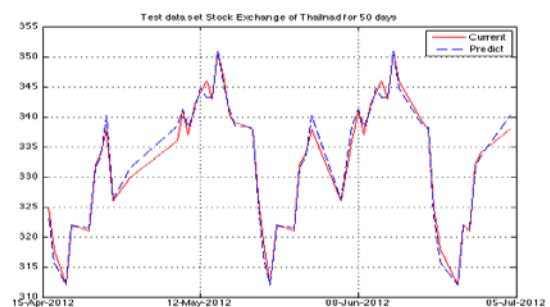
ผู้วิจัยได้ใช้ขั้นตอนวิธีแบบผสมผสานเชิงสามัญสำนึก (ITFS2HH) และใช้เครื่อง MacBook Pro Intel Core 2 Duo CPU, speed 2.66 Ghz, ram 4.00 GB, Mac Os ทำการทดลองข้อมูลที่เป็นการสอน (training set) หาจำนวน รอบ (generation) ที่เหมาะสมที่ทำให้ค่าส่งออกที่พยากรณ์ได้ตรงกับค่าที่เป็นตามความเป็นจริง ด้วยค่า RMSE ที่น้อยที่สุด พบว่าจำนวน 50 รอบให้ค่าที่เหมาะสมที่สุด โดยมีค่า RMSE=0.03982 และเมื่อนำข้อมูลการสอบ (test set) มาทำการทดสอบจะได้ค่า RMSE=0.05174 ซึ่งทำให้ได้ระบบตรรกศาสตร์แบบคลุมเครือชนิดที่ 2 ที่มีค่าพารามิเตอร์ที่เหมาะสมดังแสดงในภาพที่ 4 ในส่วนของภาพที่ 5 และ 6 แสดงตัวอย่างการ plot กราฟจำนวน 50 วันเพื่อแสดงให้เห็นว่าค่าที่พยากรณ์ได้มีค่าที่ใกล้เคียงกับค่าที่ตรงตามความเป็นจริง ภาพที่ 7 แสดง ฟังก์ชันของการเป็นสมาชิกจากการหาค่าที่เหมาะสมของ ITFS2HH และตารางที่ 2 แสดงผลของการเปรียบเทียบ ITFS2HH ในแต่ละรุ่น

ตารางที่ 2: แสดงค่าที่ได้จากการพยากรณ์ของ ITFS2HH

Method	Gen	Times/Hr.	Num of Rule	RMSE	MSE
ITFS2HH	10	6	10	0.07306	0.005337
ITFS2HH	30	12	10	0.05913	0.003496
ITFS2HH	50	24	10	0.03982	0.001585

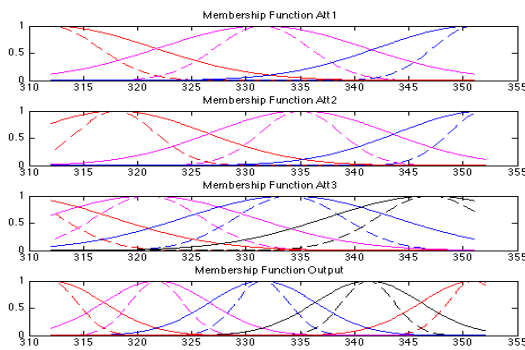


ภาพที่ 5: ITFS2HH Training set , RMSE=0.03982



ภาพที่ 6: ITFS2HH Test set , RMSE=0.05174

Method	RMSE
(1+1)ES[4]	7.44916
Meesad and Srikhacha [5]	4.04165
BPNN [6]	15.61025
ANFIS Type-I	0.02371
ITFS2HH	0.03982



ภาพที่ 7: แสดง Membership function ที่ได้จากการหาค่าที่เหมาะสมของ ITFS2HH

จากผลการทดลองที่ได้จากการพยากรณ์ด้วยเทคนิค ITFS2HH ผู้วิจัยได้ทำการเปรียบเทียบผลที่ได้กับวิธีการอื่นๆ ซึ่งผลที่ได้จากการเปรียบเทียบจะเห็นได้ว่าค่าของ RMSE น้อยกว่า หรือใกล้เคียงกับวิธีการอื่นๆ ดังแสดงในตารางที่ 3

5. บทสรุป

งานวิจัยนี้นำเสนอ ITFS2HH สำหรับการพยากรณ์อนุกรมเวลาของตลาดหลักทรัพย์แห่งประเทศไทย ผลวิจัยพบว่า ระบบให้ค่า RMSE เท่ากับ 0.03982 ถึงแม้ว่าเมื่อเปรียบเทียบกับ ANFIS Type-I แล้วจะมีค่ามากกว่าอยู่เล็กน้อย เป็นเพราะว่าค่าพารามิเตอร์ของ ANFIS มีความเหมาะสมมากกว่าเช่นค่าพารามิเตอร์ของความเป็นสมาชิก จึงเป็นแนวทางให้สามารถพัฒนาให้ดีขึ้นโดยการปรับค่าต่างๆ เช่น ค่าสมการของความถี่ ค่าพารามิเตอร์ของความเป็นสมาชิก และการกลายพันธุ์ในรูปแบบของสมการอื่น หรืออาจจะมีการหาจำนวนกฎที่เหมาะสมที่น้อยที่สุด นอกจากนี้ข้อมูลที่ใช้ในการพยากรณ์อาจจะนำเอาปัจจัยอื่นๆ เข้ามาร่วมในการพยากรณ์ เช่น ข้อมูลจากตลาดหลักทรัพย์อื่นๆ ผลกระทบทางการเมืองเป็นต้น

เอกสารอ้างอิง

[1] N. Sopipan, A. Sattayatham and S. Chongcharoen., "Forecasting Returns for the Stock Exchange of Thailand Index using Multiple Regression based on Principal

Component Analysis," *Journal of Mathematics and Statistics*, Science Publications Online, vol.9, no. 1 pp. 29-37, 2013. DOI:10.3844/jmssp.2013.29.37

[2] P. Sattayatham, N. Sopipan and B. Premanode, "Forecasting the Stock Exchange of Thailand uses Day of the Week Effect and Markov Regime Switching GRAPH," *American Journal of Economics and Business Administration*, Science Publications Online, vol. 4, no. 1, ISSN 1945-5488, pp. 84-93, 2012.

[3] P. Sutheebanjard, W. Premchaiswadi, "Forecasting the Thailand Stock Market using Evolution Strategies," *Asian Academy of Mangement Journal of Accounting and Finance*, vol. 6, no. 2, pp. 85-114, 2010.

[4] P.Sutheebanjard and W. Premchaiswadi, "Factors Analysis on Stock Exchange of Thailand (SET) Index Movement," *ICTKE2009*, Bangkok, Thailand, Dec., 2009.

[5] P. Meesad and T. Srikhacha, "Universal Data Forecasting with an Adaptive Approach and Seasonal Technique," *International Conference on Intelligent Agents & Control and Automation*, 2006.

[6] A. Esfahanipour, P. Mardani, "An AFIS Model for Stock Price Prediction: The case of Tehran Stock Exchange," *INISTA*, pp. 44-49, 2011.

[7] P. Sutheebanjard and W. Premchaiswadi, "Stock Exchange of Thailand Index Prediction using Back Propagation Neural Network," *Computer and Network Technology (ICCNT)*, pp. 377-380, 2010.

[8] H. Hostos, F. Sanabria, O. Mendez and M. Melgareio, "Towards a Coevolutionary Approach for Interval Type-2 Fuzzy Modeling," *Advances in Type-2 Fuzzy Logic Systems (T2FUZZ)*, pp. 23-30, 2011.

[9] H. Ishibuchi, T.Nakashima, and T. Murata, "Three objective genetic based machine learning for linguistic rule extraction," *Information Sciences*, vol. 136, pp. 109-133, 2001.

[10] J. Zeng and L. Wang, "A generalized model of particle swarm optimization," *Pattern Recognition and Artificial Intelligence*, vol. 18, pp. 685-688, 2005.

[11] "ราคาหุ้นวันนี้ และราคาหุ้นย้อนหลัง CSL,"

<http://www.panphol.com/data/index.php/page/stockprice>

แบบจำลองการพยากรณ์ราคาสินค้าเกษตรบนค่าเฉลี่ยเคลื่อนที่และโครงข่ายประสาทเทียม

Forecasting Agricultural Product Prices using Moving Average and Artificial Neural Network Techniques

ธัญมัย พันธุ์ชนะภากร (Tanyamai Puchanapakorn)¹ และผุสดี บุญรอด (Pudsadee Boonrawd)²

^{1,2} ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยพระจอมเกล้าพระนครเหนือ
P.Tanyamai@gmail.com¹, pudsadee@kmutnb.ac.th²

บทคัดย่อ

สินค้าเกษตรเป็นสินค้าที่มีความสำคัญต่อระบบเศรษฐกิจของประเทศไทย แต่เนื่องจากราคาสินค้ามีความคลาดเคลื่อนสูงและผันผวนบ่อย การพยากรณ์ราคาสินค้าล่วงหน้าจึงเป็นงานที่ซับซ้อน คณะวิจัยจึงได้พัฒนาแบบจำลองการพยากรณ์ราคาสินค้าเกษตรที่นำเอาวิธีการค่าเฉลี่ยเคลื่อนที่ (Moving Average) เข้ามาช่วยลดความผันผวนของราคาสินค้าเกษตรร่วมกับวิธีการโครงข่ายประสาทเทียม (Artificial Neural Networks) ในการเรียนรู้รูปแบบวัฏจักรของราคาสินค้าเกษตร โดยศึกษาจากราคาสินค้าเกษตรจำนวน 134 ชนิด ระหว่างปี พ.ศ. 2552-2556 จากการพยากรณ์พบว่าค่าความคลาดเคลื่อนระหว่างราคาสินค้าเกษตรที่เก็บได้จริง $MAPE_{(Actual)}$ กับราคาสินค้าเกษตรที่ผ่านการหาค่าเฉลี่ยเคลื่อนที่แล้ว $MAPE_{(EMA)}$ เมื่อนำไปเข้าสู่การกระบวนการเรียนรู้ด้วยโครงข่ายประสาทเทียมที่ Training Time ต่าง ๆ นั้น พบว่า $MAPE_{(EMA)}$ มีค่าดีกว่า $MAPE_{(Actual)}$ ในทุก ๆ Training Time และแบบจำลองดังกล่าวสามารถนำไปประยุกต์เพื่อพยากรณ์ราคาสินค้าเกษตรอื่น ๆ ได้ในระดับที่ดี

คำสำคัญ: ค่าเฉลี่ยเคลื่อนที่ โครงข่ายประสาทเทียม ราคา
สินค้าเกษตร

Abstract

Agricultural products are vitally important to the economy of Thailand. However, the prices of agricultural products are highly affected by deviation and fluctuation. Therefore, forecasting agricultural prices can be a very complicated process. The researcher of this paper has developed models with the objective of accurately

forecasting agricultural prices. Moving Average (MA) and Artificial Neural Network (ANN) techniques were used to reduce the fluctuation of Agricultural price based on 134 agricultural products during 2009-2013. From the analyzed data, it was found that $MAPE_{(EMA)}$, price from Moving Average technique was better than $MAPE_{(Actual)}$ deviation value from the forecast when applied to the learning process of ANN in each training test. In addition, this developed model can also be applied to forecast agricultural product price in higher levels.

Keywords: Moving Average, Artificial Neural Network, Agricultural Product Prices.

1. บทนำ

สินค้าเกษตรเป็นสินค้าที่มีความสำคัญอย่างมากในครัวเรือนและอุตสาหกรรม [1] เนื่องจากราคาสินค้าเกษตรจัดว่าเป็นปัจจัยหนึ่งที่ต้องใช้เพื่อการดำรงชีวิต การซื้อสินค้าเกษตรเพื่อใช้ภายในครัวเรือนอาจไม่เห็นถึงผลกระทบของการขึ้นลงของราคาสินค้าเกษตรมากนัก แต่หากจำเป็นต้องซื้อสินค้าเกษตรจำนวนมากหรือยกลิตรัมต่อวันอย่างเช่น โรงงานอุตสาหกรรมอาหารขนาดใหญ่แล้ว ทุก ๆ บาทที่ราคาเพิ่มขึ้นนั้นหมายถึงต้นทุนที่โรงงานต้องจ่ายมากขึ้น เพื่อผลิตสินค้าในปริมาณเท่าเดิม เนื่องจากโรงงานอุตสาหกรรมไม่สามารถกำหนดราคากลางของสินค้าเกษตรได้ แต่ถ้าหากโรงงานอุตสาหกรรมสามารถทราบราคาของสินค้าเกษตรที่ต้องการซื้อได้ล่วงหน้าได้ ก็จะสามารถประมาณการต้นทุนของสินค้าที่จะทำการผลิต วางแผนการผลิต จัดสรรกำลังคน จัดการสินค้าคงคลังได้ ทำให้เกิดการบริหารจัดการที่มีประสิทธิภาพมากขึ้น นำมาซึ่ง

การลดต้นทุนในอีกทางหนึ่ง แต่การได้มาซึ่งราคาสินค้าเกษตร ต้องอาศัยปัจจัยแวดล้อมทางธุรกิจหลากหลายประการ เช่น ความสามารถในการต่อรองราคา ความต้องการของตลาด ปัญหาอุทกภัย ภัยแล้ง ภาวะคู่แข่งทางตลาด ตลอดจนต้นทุนของการผลิตสินค้าเกษตร เป็นต้น ปัจจัยเหล่านี้ล้วนส่งผลกระทบต่อถึงราคาสินค้าเกษตรทั้งสิ้น ปัจจัยบางตัวอาจส่งผลให้ราคาสินค้าเกษตรเพิ่มขึ้นเป็นเท่าตัว บางตัวอาจมีกระทบเพียงเล็กน้อย ดังนั้นการคาดการณ์ถึงราคาสินค้าเกษตรจึงเป็นเรื่องที่ต้องอาศัยความชำนาญ ประสบการณ์ สถิติข้อมูลเก่า (History data) เพื่อนำเอาข้อมูลที่เก็บรวบรวมไว้มาช่วยในการพยากรณ์ราคาสินค้าเกษตรที่จะเปลี่ยนแปลงไปตามสถานการณ์ต่าง ๆ โดยแนวทางที่นำเสนอต้องขึ้นอยู่กับพื้นฐานของความน่าเชื่อถือ มีหลักการที่แน่นอนเป็นเหตุเป็นผล ใช้งานได้ง่าย ไม่ซับซ้อนเกินไปต่อการทำความเข้าใจ และยืดหยุ่นต่อการแก้ไขปัญหาในสถานการณ์ต่าง ๆ

วิธีการจำนวนมากถูกนำมาประยุกต์ใช้ในการพยากรณ์ข้อมูลร่วมกับหลักการต่าง ๆ [2], [3] เช่น การใช้ค่าเฉลี่ยเคลื่อนที่ในการพยากรณ์ราคาหุ้น เนื่องจากราคาของตลาดหุ้นมีการแกว่งตัวและผันผวนสูง ทำให้การพิจารณาเส้นแนวโน้มทำได้ยาก วิธีการดังกล่าวจึงถูกนำมาเพื่อแก้ค่าสถิติของข้อมูลให้เห็นเส้นแนวโน้มที่ชัดเจนขึ้น ซึ่งขณะวิจัยสังเกตเห็นว่าวิธีการดังกล่าวสามารถนำมาประยุกต์ใช้กับหลักการพยากรณ์ราคาสินค้าเกษตรได้เช่นกัน เพื่อช่วยลดการแกว่งตัวของราคาสินค้า และทำการกำจัดค่าที่อาจเกิดจากการเก็บข้อมูลที่ผิดพลาดได้ สามารถมองเห็นเส้นแนวโน้มของราคาสินค้าได้ง่ายขึ้น เมื่อนำมาใช้ร่วมกับโครงข่ายประสาทเทียมในการเรียนรู้ จดจำรูปแบบการขึ้นลงของราคาสินค้าทำให้การพยากรณ์ราคาสินค้าเกษตรมีความแม่นยำมากขึ้น ซึ่งจะกล่าวถึงต่อไปในบทที่ 2 ทฤษฎีที่เกี่ยวข้องกับงานวิจัย วิธีการดำเนินงานวิจัยในบทที่ 3 สรุปผลการวัดประสิทธิภาพในบทที่ 4 และสรุปผลในบทที่ 5

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การพยากรณ์ข้อมูลอนุกรมเวลา (Time Series Prediction)

เป็นการคาดการณ์ข้อมูลหรือค่าสังเกตที่เปลี่ยนแปลงไปตามลำดับเวลาที่เกิดขึ้นอย่างต่อเนื่องโดยข้อมูลเหล่านี้ถูกเก็บรวบรวม ณ ช่วงเวลาต่าง ๆ เช่นรายวันรายสัปดาห์รายเดือนราย

ไตรมาสหรือรายปีซึ่งลักษณะข้อมูลที่จะนำมาวิเคราะห์อนุกรมเวลาได้นั้นต้องมีคุณสมบัติ ดังนี้

1. เป็นข้อมูลที่เก็บรวบรวมเป็นระยะเวลาานพอสมควร
2. เป็นข้อมูลที่มีความสัมพันธ์กับเวลา
3. ระยะเวลาที่ใช้ในการเก็บรวบรวมข้อมูลจะต้องเท่ากัน

2.2 การหาค่าเฉลี่ยเคลื่อนที่ (Exponential Moving Average: EMA)

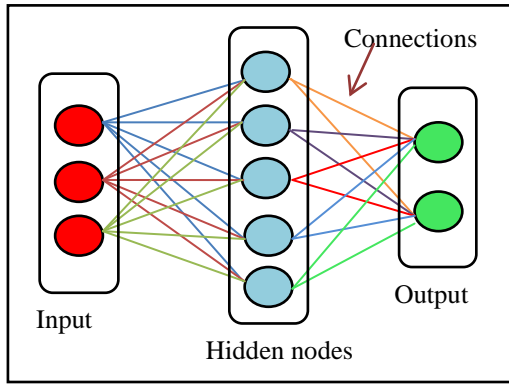
เป็นค่าเฉลี่ยที่เกิดจากการนำเอาข้อมูลช่วงที่สนใจนำมาทำการหาค่าเฉลี่ย เพื่อเป็นการเฉลี่ยค่าของข้อมูลที่ได้รับการผิดพลาด หรือมีความผันผวนมาก ให้มีความผันผวนน้อยลง ซึ่งเมื่อมีการเฉลี่ยค่าของข้อมูลแล้ว จะทำให้เห็นแนวโน้มของชุดข้อมูลได้ง่ายขึ้น การคำนวณค่าเฉลี่ยเคลื่อนที่มีหลายวิธีการ เช่น Simple Moving Average: SMA, Exponential Moving Average: EMA และ Time series Moving Average: TMA เป็นต้น

งานวิจัยนี้ได้นำเอาวิธีการ SMA และ EMA เข้ามาช่วยในการเฉลี่ยค่าข้อมูลที่มีความผันผวน จากการศึกษางานวิจัยที่เกี่ยวข้องพบว่า การใช้ค่าเฉลี่ยเคลื่อนที่แบบเอกซ์โพเนนเชียลที่ให้ค่าสัดส่วนความถูกต้องมากที่สุด คือ การคำนวณจากค่า EMA ที่ 75 วัน [4], [5] ซึ่งจะเป็นค่า EMA ที่ถูกนำมาใช้ในการคำนวณในงานวิจัยนี้เช่นกัน โดยที่ค่า EMA₇₅ จะถูกคำนวณได้ต้องผ่านการคำนวณหาค่า SMA₇₅ เสียก่อน

2.3 วิธีการโครงข่ายประสาทเทียม (Artificial Neural Network: ANN)

เป็นแบบจำลองทางคณิตศาสตร์ที่ถูกสร้างขึ้นมาเพื่อจำลองการทำงานให้คล้ายคลึงกับการทำงานของระบบประสาทในสมองของมนุษย์ โดยมีวัตถุประสงค์เพื่อให้ระบบสามารถรู้จำรูปแบบ (Pattern Recognition) [6] และนำมาสร้างองค์ความรู้ใหม่ได้ (Knowledge Extraction) เป็นต้น

โครงข่ายประสาทเทียม จะประกอบไปด้วยอินพุตโหนด (Input nodes) และเอาต์พุตโหนด (Output nodes) โดยแต่ละโหนดจะมีการเชื่อมโยงกันเหมือนเครือข่ายประสาท มีการกำหนดค่าน้ำหนัก (Weight) ที่แตกต่างกันออกไป และเมื่อข้อมูลเข้ามาที่อินพุตโหนดและนำมาคูณกับค่าน้ำหนักที่กำหนดไว้ ในเบื้องต้น ผลลัพธ์ที่ได้จากทุก ๆ ขาจะถูกนำมารวมกันเทียบกับค่าที่กำหนดไว้ (Threshold) [7], [8] เมื่อค่าผลลัพธ์ที่ได้มากกว่าค่าที่กำหนดไว้ ระบบก็จะส่งเอาต์พุตออกไป แต่ถ้าน้อยกว่า ระบบก็จะไม่ส่งเอาต์พุตออกไปดังแสดงในภาพที่ 1



ภาพที่ 1: โครงสร้างเครื่องข่ายประสาทเทียม

โดยมีหลักการเรียนรู้ 2 รูปแบบ คือ

ก) การเรียนรู้แบบมีการสอน (Supervised Learning) คือ การพิจารณาจากคำตอบที่ระบบให้ผลลัพธ์ออกมาเทียบกับข้อมูลที่เกิดขึ้นจริง ระบบจะเรียนรู้ในการปรับค่าที่คลาดเคลื่อนไปให้ใกล้เคียงกับค่าที่เกิดขึ้นจริง

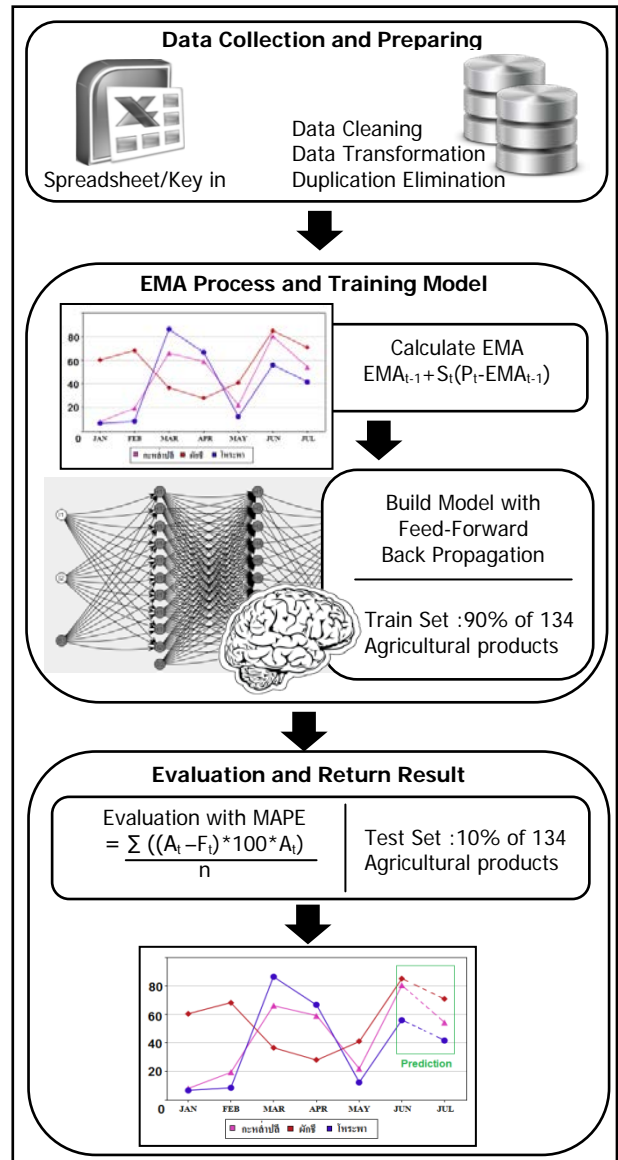
ข) การเรียนแบบไม่มีการสอน (Unsupervised Learning) คือ การเรียนรู้ในการแบ่งแยกกลุ่มข้อมูลตามลักษณะของข้อมูล ซึ่งระบบสามารถจัดหมวดหมู่ของข้อมูลได้ เช่น ระบบสามารถแบ่งแยกระหว่างพืชและสัตว์ได้

2.4 วรรณกรรมที่เกี่ยวข้อง

นักวิจัยจำนวนมากได้ทำการศึกษาและพัฒนา วิธีการ กรอบแนวคิดต่าง ๆ ที่เกี่ยวข้องกับการพยากรณ์ข้อมูล โดยรวบรวมทฤษฎีต่าง ๆ ที่ถูกกล่าวอ้างในหลาย ๆ งานวิจัยมาเพื่อหาข้อสรุปถึงข้อดีข้อเสียของแต่ละทฤษฎีที่ถูกนำมาใช้ เช่น วิธีการค่าเฉลี่ยเคลื่อนที่ การพยากรณ์อนุกรมเวลาด้วยตัวแบบ ARIMA วิธีการบอซ-เจนกินส์ วิธีการพยากรณ์ของโฮลท์ วิธีการพยากรณ์ของการ์ดเนอร์-แมคเคนซีและวิธีการพยากรณ์รวม เป็นต้น ซึ่งตอนหนึ่งของหัวข้อ Neural Net ได้มีกล่าวอ้างว่า [2], [7], [9] การพยากรณ์ข้อมูลด้วยวิธีการโครงข่ายประสาทเทียมนั้น หากข้อมูลที่ต้องการศึกษามีรูปแบบที่เป็นข้อมูลอนุกรมเวลา อย่างเช่น การพยากรณ์ราคาหุ้น การพยากรณ์ผู้ป่วยโรคเบาหวาน การพยากรณ์ปริมาณการใช้งานไฟฟ้าในครัวเรือน เป็นต้น หากมีการกำหนดอินพุต และแบบจำลองได้เรียนรู้ด้วยข้อมูลในปริมาณมากพอ จะยิ่งทำให้ได้ผลลัพธ์ของการพยากรณ์ที่ดี และมีประสิทธิภาพกว่าการใช้งานวิธีการอื่น ๆ เนื่องจากระบบได้มีการเรียนรู้จากข้อมูลเก่ามาก่อนแล้ว [10]

3. วิธีดำเนินงานวิจัย

การพยากรณ์ข้อมูลราคาสินค้าเกษตรโดยใช้วิธีการการหาค่าเฉลี่ยเคลื่อนที่แบบเอกซ์โพเนนเชียลและวิธีการโครงข่ายประสาทเทียม คณะวิจัยได้กำหนดขั้นตอนการดำเนินการทั้งหมด 3 ขั้นตอน ซึ่งแสดงในกรอบแนวคิดภาพรวมระบบในภาพที่ 2



ภาพที่ 2: กรอบแนวคิดภาพรวมระบบ

3.1 การรวบรวมและวิเคราะห์ข้อมูล (Data Collection)

ข้อมูลราคาสินค้าเกษตรย้อนหลังถูกรวบรวมจาก บริษัทไทยแอ็กโกรเอ็กซ์เชนจ์ จำกัด และบริษัท ดอนเมืองพัฒนา จำกัด ซึ่งเผยแพร่ให้แก่ประชาชนทั่วไปสามารถนำไปใช้ประโยชน์ได้ โดยลักษณะข้อมูลถูกจัดเก็บในรูปแบบข้อมูลอนุกรมเวลาหรือข้อมูลที่ถูกรวบรวมไว้ในแต่ละคาบเวลา มี

ความถี่ในการบันทึกข้อมูลเป็นรายวันย้อนหลัง ตั้งแต่ มกราคม พ.ศ. 2552 – ธันวาคม พ.ศ. 2556 มีจำนวนทั้งสิ้น 243 ชนิด จำนวน 544,265 เรคอร์ด มีแอตทริบิวต์ทั้งหมด 7 แอตทริบิวต์ คือ ชื่อสินค้า วันที่เก็บราคา หน่วยขาย ราคาขายต่อหน่วย เบอร์สินค้า เกรดสินค้า และกลุ่มสินค้า โดยมี 4 Classes คือ กลุ่มผักทั่วไป กลุ่มผักปรุงรส กลุ่มผักเมืองหนาว และกลุ่มผักพื้นบ้าน ซึ่งรายละเอียดของชุดข้อมูลมีไฟล์ที่สำคัญ ดังนี้

ตารางที่ 1: รายละเอียดของชุดข้อมูลสินค้าเกษตรที่เก็บรวบรวม

ชื่อฟิลด์	ความหมาย
Product Name	ชื่อสินค้าเกษตร
Product Group	กลุ่มสินค้าเกษตร
Price	ราคาเฉลี่ยที่ทำการเก็บได้
Unit	หน่วยซื้อขาย
Size	ขนาดสินค้า
Grade	เกรดสินค้า
Date	วันที่ทำการเก็บราคา

3.2 การหาค่าเฉลี่ยเคลื่อนที่และเรียนรู้แบบจำลอง (EMA Process and Training Model)

การเตรียมข้อมูลเพื่อใช้สำหรับการทำดาต้าไมนิ่ง (Data Preparation) เริ่มต้นจากนำข้อมูลที่ได้จากทั้งสองแหล่งข้อมูลมารวมกัน (Data Integration) และทำการตัดรายการสินค้า ตัดแอตทริบิวต์ที่ไม่เกี่ยวข้อง เพื่อให้เวลาในการประมวลผลข้อมูลลดลง และมีความถูกต้องมากยิ่งขึ้น (Data Cleaning) เนื่องจากข้อมูลที่นำมาใช้มีโอกาสเกิดค่าสูญหาย (Missing value) สูงมากจากหลาย ๆ ปัจจัย เช่น ช่วงเวลาสินค้าขาดตลาด ความผิดพลาดจากการบันทึกข้อมูล และสินค้าในตลาดไม่สามารถเป็นตัวแทนราคาของสินค้าในวันนั้นได้ คณะวิจัยจึงได้ทำการจัดเรียงเรคอร์ดที่หายไป เพิ่มแอตทริบิวต์บ่งชี้เรคอร์ด Missing Value และแอตทริบิวต์วันหยุดเทศกาลในแต่ละปี ปฏิทินซึ่งเป็นปัจจัยหนึ่งที่มีผลต่อการขึ้นลงของราคาสินค้าเกษตร (Data Transformation) จากนั้นแทนที่ค่า Missing value ด้วยค่า 2 ประเภท คือ

ค่า 0 แทนด้วย Price₀

ค่าเฉลี่ยราคาในปีนั้น ๆ แทนด้วย Price_{AVG}

จากนั้นนำราคาทั้งสองมาคำนวณหาค่าเฉลี่ยเคลื่อนที่อย่างง่าย SMA₇₅ โดยที่

$$SMA = \frac{(P_t + P_{t-1} + P_{t-2} + \dots + P_{t-n+1})}{n} \quad (1)$$

เมื่อ n แทนค่า จำนวนวันของค่าเฉลี่ย, P_t แทนค่า ราคาที่เลือกใช้ในการคำนวณ (เช่น ราคาสินค้าเฉลี่ย) ณ วันปัจจุบัน และ P_{t-n} แทนค่า ราคาที่เลือกใช้ในการคำนวณย้อนหลังไปจำนวน n คาบ คณะวิจัยแทน Price₀ ที่ผ่านกระบวนการหา SMA ว่า SMA₀ และแทน Price_{AVG} ที่ผ่านกระบวนการหา SMA ว่า SMA_{AVG} เมื่อได้ราคาสินค้าที่ผ่านขั้นตอนการหาค่าเฉลี่ยเคลื่อนที่อย่างง่ายแล้วนั้น จึงนำไปหาค่าเฉลี่ยเคลื่อนที่เอกซ์โพเนนเชียลที่ 75 วัน (EMA₇₅) ดังสมการที่ 2

$$EMA_t = EMA_{t-1} + S_f (P_t - EMA_{t-1}) \quad (2)$$

เมื่อ EMA_t แทนค่า Exponential Moving Average ณ เวลาปัจจุบัน, EMA_{t-1} แทนค่า Exponential Moving Average ณ คาบเวลาก่อนหน้า, S_f แทนค่า Smoothing Factor เท่ากับ $\frac{2}{n+1}$, P_t แทนราคาปัจจุบัน และ n แทนจำนวนวัน ซึ่งรหัสจำลองของ EMA ที่นำมาใช้งานวิจัยนี้ แสดงดังภาพที่ 3

```

Class Calculate_EMA {
    Private Double SmoothF;
    Private Double OldPrice;
    Public EMA (Double SmoothF) {
        This.SmoothF = SmoothF;
    }
    Public Double Average (Double Price) {
        If (OldPrice==NULL) {
            OldPrice = Price;
            Return Price;
        }
        Double NewPrice = OldPrice + SmoothF * (Price-OldPrice);
        OldPrice = NewPrice;
        Return NewPrice;
    }
}
    
```

ภาพที่ 3: รหัสจำลองการคำนวณ EMA

โดยที่ Smooth F คือ ค่า Smoothing Factor

OldPrice คือ ค่า EMA ณ คาบเวลาก่อนหน้า

Price คือ ราคา ณ ปัจจุบัน

แทนสัญลักษณ์ SMA₀ ที่ผ่านกระบวนการหา EMA ว่า EMA₀ และแทน SMA_{AVG} ที่ผ่านกระบวนการหา EMA ว่า EMA_{AVG}

การสร้างแบบจำลองและการสอนข้อมูล คณะวิจัยได้พัฒนาแบบจำลองด้วยโปรแกรมการทำเหมืองข้อมูลเวลา (WEKA) รุ่น 3.7.10 ซึ่งเป็นโปรแกรมสำเร็จรูปถูกพัฒนาขึ้นด้วยภาษาจาวา คณะวิจัยได้ทำการทวนสอบการแบ่งกลุ่มของสินค้าจากข้อมูลที่ได้รับมาด้วยวิธีการ SimpleKmeans จากนั้นจึงนำไปหาแบบจำลองด้วยโครงข่ายประสาทเทียมแบบเปอร์

เซียบตรอนหลายชั้น (Multilayered Feed-Forward Neural Network) ร่วมกับการแพร่กระจายย้อนกลับ (Back propagation) ซึ่งเป็นโครงข่ายประสาทเทียมที่ข้อมูลที่ถูกประมวลผลจะถูกส่งจากอินพุตโหนดไปยังโหนดข้างหน้าเรื่อย ๆ ในทิศทางเดียวจนถึงเอาต์พุตโหนดโดยไม่ย้อนกลับแม้แต่ในเลเยอร์เดียวกัน มีเพียงค่าน้ำหนัก Weight เท่านั้นที่จะถูกส่งกลับ ไปเพื่อทำการปรับปรุ้ค่าน้ำหนักที่ได้หลังจากกระบวนการเรียนรู้แล้ว โดยทำการกำหนด Training Set ที่ 90% และใช้ Test Set ที่ 10% ผลลัพธ์ที่ได้จากการเรียนรู้แบบจำลองคือ ราคาสินค้าเกษตรที่ถูกพยากรณ์ล่วงหน้า 5 วัน และค่า MAPE ที่ได้หลังจากการพยากรณ์ ณ Training Time ต่าง ๆ

3.3 การวัดประสิทธิภาพระบบ (Evaluation and Return Result)

การทดสอบแบบจำลองการพยากรณ์จะนำเอาค่าความผิดพลาดร้อยละเฉลี่ยสัมบูรณ์ (MAPE) มาใช้ในการเปรียบเทียบระหว่างการพยากรณ์ข้อมูลราคาสินค้า Price_{Actual} และ Price_{EMA} ว่าการพยากรณ์วิธีการใดจะมีค่าความผิดพลาดร้อยละเฉลี่ยสัมบูรณ์น้อยที่สุด โดยคำนวณได้จากผลรวมของส่วนต่างระหว่างค่าจริงกับค่าที่พยากรณ์ได้มาคูณกับค่าจริงคูณ 100 และหารด้วยกลุ่มตัวอย่างทั้งหมด สามารถเขียนได้ดังสมการที่ 3

$$MAPE = \frac{\sum ((A_t - F_t) * 100 * A_t)}{n} \quad (3)$$

เมื่อ A_t คือ ค่าที่เกิดขึ้นจริง ณ เวลาที่ใด ๆ

F_t คือ ค่าที่พยากรณ์ ณ เวลาที่ใด ๆ

n คือ จำนวนกลุ่มตัวอย่างทั้งหมด

ซึ่งค่า MAPE ยิ่งน้อย หมายถึง การพยากรณ์ยิ่งมีความแม่นยำมาก

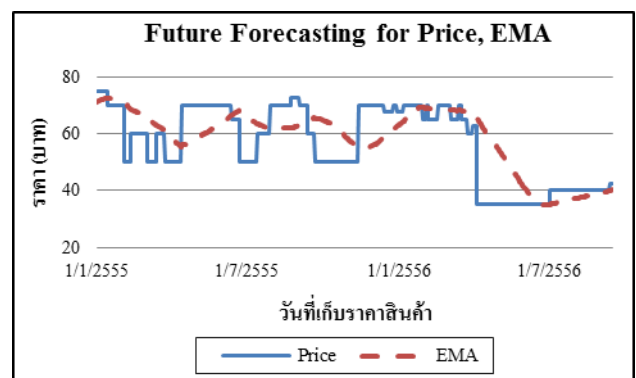
4. ผลการวัดประสิทธิภาพ

จากการการแบ่งกลุ่มของสินค้าเกษตรจำนวน 134 ชนิดด้วยวิธีการ SimpleKMeans พบว่าสินค้าถูกแบ่งออกเป็นทั้งสิ้น 3 กลุ่ม คือ กลุ่มที่ 1 ประกอบไปด้วยกลุ่มผักทั่วไปและผักพื้นบ้าน กลุ่มที่ 2 คือ ผักปรุ้รส และกลุ่มสุดท้าย คือ ผักเมืองหนาว ดังแสดงตัวอย่างในตารางที่ 2

ตารางที่ 2 : ผลลัพธ์การแบ่งกลุ่มด้วยวิธีการ SimpleKMeans

กลุ่มสินค้า	สินค้าเกษตร	กลุ่ม			
		1	2	3	4
พืชปรุ้รส	กะเทียมจีนใหญ่	0	1	2,207	1
พืชปรุ้รส	กะเทียมปิงปอง	0	1	2,214	0
พืชฤดูหนาว	ผักโขม	6	0	1	3,298
พืชฤดูหนาว	พื้กทองญี่ปุ่น	1	0	0	2,225
พืชทั่วไป	พื้กทองเกษตรใหญ่	0	2,575	0	0
พืชทั่วไป	พื้กทองเกษตรกลาง	0	2,575	0	0
พืชพื้นบ้าน	ผักชี	0	3,306	0	0
พืชพื้นบ้าน	ยอดพื้กทอง	1	2,267	1	1

เมื่อนำสินค้าทั้ง 3 กลุ่มเข้าสู่กระบวนการเรียนรู้โดยกำหนดอินพุตโหนด 7 โหนด จำนวนชั้นซ่อน a ชั้น และเอาต์พุตโหนด 4 โหนด คือ ราคาที่เก็บได้จริงแทนค่า Missing value ด้วย 0 Price₀ ราคาที่เก็บได้จริงแทนค่า Missing value ด้วยราคาเฉลี่ยสินค้าในปีนั้น ๆ คือ Price_{AVG} ราคาที่ผ่านการหา Moving Average แทนค่า Missing value ด้วย 0 คือ EMA₀ และราคาที่ผ่านการหา Moving Average แทนค่า Missing value ด้วยราคาเฉลี่ยสินค้าในปีนั้น ๆ คือ EMA_{AVG} โดยโครงสร้างของ ANN ในงานวิจัยนี้คือ 7-a-4 โดยใช้พารามิเตอร์ค่า Learning Rate และค่า Momentum ตามค่าดีฟอลต์คือ 0.3 และ 0.2 ตามลำดับ และค่า Training Time ที่ 300 500 800 1,000 1,500 และ 2,000 รอบ ซึ่งผลการพยากรณ์คือราคาสินค้าเกษตรล่วงหน้า 5 วัน จะถูกแสดงผลออกในรูปแบบกราฟ ดังแสดงตัวอย่างในภาพที่ 4



ภาพที่ 4: ตัวอย่างการพยากรณ์ราคาสินค้าเกษตรล่วงหน้า ซึ่งผลการประเมินประสิทธิภาพการพยากรณ์ด้วยค่า MAPE ถูกบันทึกแยกตามกลุ่มสินค้า รายการสินค้า และ

ประเภทของราคาที่ใช้พยากรณ์ตามจำนวนรอบ Training Time ต่าง ๆ ดังแสดงตัวอย่างในตารางที่ 3

ตารางที่ 3: ตัวอย่างค่า MAPE ที่ได้หลังจากการพยากรณ์ราคาสินค้าเกษตรล่วงหน้า 5 วัน กลุ่มผักทั่วไป

สินค้าเกษตร		300 รอบ		500 รอบ	
		Actual	EMA	Actual	EMA
ผัก ทั่วไป	เห็ดเข็มทอง	31.53	1.10	36.66	1.38
	เห็ดหอมสด	20.25	1.70	18.69	1.67
	กระเจี๊ยบเขียว	12.14	5.40	13.82	5.39
	มะเขือสะตั้ง	14.16	2.46	14.82	2.24
	เห็ดนางฟ้า	32.80	2.53	26.67	2.66
	เห็ดนางรม	104.58	7.94	123.25	7.59
	เห็ดเป๋าฮื้อ	120.22	5.53	131.15	6.25
	เห็ดฟาง	9.94	0.88	10.23	0.93
	เห็ดหูหนู	20.22	1.64	20.56	1.37

5. สรุป

จากตารางผลการดำเนินงานที่ 3 ทำให้สามารถสรุปได้ว่าการพยากรณ์ราคาสินค้าเกษตรล่วงหน้า 5 วัน ที่ค่า Learning Rate 0.3 และค่า Momentum 0.2 พบว่า Training Time ที่ให้ค่าการพยากรณ์แม่นยำที่สุด หรือมีค่าความผิดพลาดร้อยละเฉลี่ยสัมบูรณ์ (MAPE) ดีที่สุด คือ 300 รอบ

เมื่อเปรียบเทียบวิธีการแทนค่า Missing value ระหว่าง $Price_0$, $Price_{AVG}$, EMA_0 และ EMA_{AVG} พบว่าค่า MAPE จากการพยากรณ์ค่า $Price_{AVG}$ จะดีกว่า $Price_0$ มากกว่า 78% ของชนิดสินค้า แต่ค่า MAPE จากการพยากรณ์ค่า EMA_0 กับ EMA_{AVG} ยังไม่สามารถบ่งชี้ได้ชัดเจนว่าวิธีการแทนค่าทั้ง 2 วิธี ใดดีกว่า เนื่องจากสัดส่วนของ MAPE ที่ดีที่สุดระหว่าง EMA_0 และ EMA_{AVG} มีค่าใกล้เคียงกัน ขึ้นอยู่กับแต่ละชนิดสินค้า

แต่ทั้งนี้ค่าการพยากรณ์ที่ได้จาก $Price_{EMA}$ มีค่า MAPE ต่ำกว่าค่า $Price_{Actual}$ เสมอ แม้ว่าค่าการพยากรณ์ในบางสินค้าจะมีค่า MAPE $_{Actual}$ สูงมากผิดปกติเนื่องจากค่า Missing value แต่เมื่อผ่านกระบวนการหาค่าเฉลี่ยเคลื่อนที่ก่อนนำไปเรียนรู้ในโครงข่ายประสาทเทียมทำให้ค่า $MAPE_{EMA}$ ลดลงอย่างมาก นั้นหมายความว่า การพยากรณ์โดยใช้วิธีการค่าเฉลี่ยเคลื่อนที่ร่วมกับวิธีการโครงข่ายประสาทเทียมเพื่อพยากรณ์ข้อมูลอนุกรมเวลาของราคาสินค้าเกษตรนั้นดีกว่าการเลือกใช้วิธีการพยากรณ์ด้วยวิธีการใดวิธีการหนึ่ง และเพื่อให้ค่าความ

ผิดพลาดร้อยละเฉลี่ยสัมบูรณ์ที่ดีที่สุด แบบจำลองควรมีการปรับค่าพารามิเตอร์ของ Learning Rate, Momentum และจำนวนชั้นซ่อน ภายในโครงข่ายประสาทเทียมต่อไป เพื่อให้ได้แบบจำลองการพยากรณ์ที่มีประสิทธิภาพมากขึ้น

เอกสารอ้างอิง

- [1] ลดาวัลย์ คำภา, “นโยบายการพัฒนาการเกษตรและสินค้าเกษตรไทย เพื่อเพิ่มขีดความสามารถในการแข่งขัน” *การประชุมวิชาการวิทยาการหลังการเก็บเกี่ยว คณะทรัพยากรชีวภาพและเทคโนโลยี มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี*, 28 มิถุนายน 2550
- [2] Gooijer Jan G. De, Hyndman J Rob, “25 years of time series forecasting”. *International Journal of Forecasting*, Vol. 22, pp. 443–473, 2006
- [3] ยูพิน กาญจนะศักดิ์ดา และคณะ, “การเปรียบเทียบวิธีพยากรณ์ราคาทองคำแท่งรายวันระหว่างวิธีบอซซ์-เจโนกินส์ วิธีการพยากรณ์ของโสมท์ วิธีการพยากรณ์ของการ์ดเนอร์-แมกเคนซี และวิธีการพยากรณ์รวม”. *ผลงานวิจัยและวิชาการด้านวิทยาศาสตร์และเทคโนโลยี คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยหอการค้าไทย*.
- [4] ชื่นใจ สุกปาน, “แบบจำลองเพื่อการพยากรณ์ราคาผลผลิตทางการเกษตรในตลาดการซื้อขายล่วงหน้า: กรณีศึกษายางพารา”. *การประชุมวิชาการสถิติประยุกต์และเทคโนโลยีสารสนเทศระดับชาติ ปีที่ 2556* หน้า 23-37.
- [5] ชวริย์ รัตนวงศ์ศรี, “การพยากรณ์ราคาของพาราในตลาดสินค้าเกษตรล่วงหน้าแห่งประเทศไทย” *Veridian E-Journal*, SU, Vol. 6, No. 1 January-April, 2013
- [6] สันติภาพ พันธุ์เสือ และมณฑิเร รัตนศิริวงศ์วุฒิ, “ระบบการพยากรณ์ประเภทผู้ป่วยโรคเบาหวานโดยใช้โครงข่ายประสาทเทียม.” *การประชุมวิชาการเสนอผลงานวิจัยระดับบัณฑิตศึกษา แห่งชาติครั้งที่ 29* ปีที่ 2556 หน้า 648-655.
- [7] Tsung-Sheng Chang, “A comparative study of artificial neural networks, and decision trees for digital game content stocks price prediction”. *Expert Systems with Applications*, Vol.38, pp.14846–14851, 2011.
- [8] Coskun Hamzacebi, Diyar Akay, Fevzi Kutay, “Comparison of direct and iterative artificial neural network forecast approaches in multi-periodic time series forecasting”. *Expert Systems with Applications*, Vol. 36, pp. 3839–3844, 2009.
- [9] Farshid Keynia, “A new feature selection algorithm and composite neural network for electricity price forecasting”. *Engineering Applications of Artificial Intelligence*, Vol. 25, pp. 1687–1697, 2012.

การประยุกต์ใช้วิธีวิเคราะห์ความถดถอยเชิงเส้นสำหรับประเมินพื้นที่ใบพืช

Apply linear regression analysis to estimate leaf area.

สัตตภาภูมิ ไทยพานิช¹(Sattarpoom Thaiparnit) บารมี โอสธีรกุล¹(Baramee Osateerakul)
และ มหศักดิ์ เกตุฉ่ำ²(Mahasak Ketcham)

¹คณะบริหารธุรกิจและเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ

²คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
sattarpoom@hotmail.com, b_eagle9@hotmail.com, maoqueue@hotmail.com

บทคัดย่อ

การวัดพื้นที่ของใบพืชเป็นสิ่งจำเป็นในการวิเคราะห์อัตรา การสังเคราะห์ด้วยแสงของพืชบทความนี้นำเสนอการสร้าง แบบจำลองสำหรับประเมินพื้นที่ใบพืชโดยประยุกต์ใช้วิธี วิเคราะห์ความถดถอยเชิงเส้นในการประเมินพื้นที่จากค่า จำนวนจุดสีของใบพืช ในการสร้างแบบจำลองนั้นกระดาษ ขนาดมาตรฐานจำนวน 100 ขนาดที่มีความแตกต่างกันถูก สแกนด้วยเครื่องสแกนเนอร์ที่ความละเอียด 100 และ 200 พิกเซล ถูกนำมาวิเคราะห์ด้วยวิธีวิเคราะห์ความถดถอยเพื่อ ได้ แบบจำลองสำหรับประเมินพื้นที่ใบพืช และ ใบพืชจริงที่มี รูปร่างและขนาดที่ต่างกันจำนวน 60 ใบ ถูกนำมาใช้ ทดสอบหาค่าความคลาดเคลื่อนของแบบจำลอง ผลการทดลอง พบว่าแบบจำลองทั้งสามมีความคลาดเคลื่อนต่ำ และพบว่า การ สแกนที่ความละเอียดที่สูงขึ้นจะส่งผลต่อความคลาดเคลื่อนที่ ลดลง

คำสำคัญ: การวิเคราะห์ความถดถอย, การประเมินพื้นที่ใบ

Abstract

Measurements of leaf area is essential to analyze the rate of photosynthesis of plants, this article presents a modeling for assessing plant leaf area by applying linear regression analysis to evaluate hosting packages. Rather the number of dots of plant leaves. In modeling the standard sizes of 100 sizes with different resolution scans were 100 and 200 pixels. And actually shaped leaves and 60 leaves of different size were used to test the accuracy of the model. The results showed that both models are offered with low tolerances.

Keyword: Regression analysis ,Leaf area estimation

1. บทนำ

กระบวนการสังเคราะห์ด้วยแสงนับว่ามีความสำคัญและเป็นปัจจัยที่มีความสำคัญต่อกระบวนการผลิตอาหารของพืชทุกชนิด ขนาดพื้นที่ของใบพืชย่อมส่งผลต่อความสามารถในการสังเคราะห์ด้วยแสงของพืช[1] บริเวณใบมีพื้นที่ทางฝั่งด้านหนึ่งของใบจะเป็นสีเขียว ซึ่งแสดงถึงความสามารถที่จะทำการสังเคราะห์แสงยังคงอยู่ การวัดพื้นที่ของใบพืชนั้นสามารถทำได้หลายวิธีด้วยกันเช่น [2]

- 1) การวัดขนาดพื้นที่ใบด้วยแบบจำลองความสัมพันธ์ระหว่างความยาวสูงสุดและความกว้างสูงสุดของใบพืช
- 2) การวัดขนาดพื้นที่ใบด้วยวิธีการนับตารางจากกระดาษตารางมาตรฐาน
- 3) การประเมินขนาดพื้นที่ใบด้วยเทคนิคการทำให้ตกตะกอน
- 4) การประเมินขนาดพื้นที่ใบด้วยการเจาะใบเพื่อชั่งน้ำหนัก

ความยากในการ วัดและความผิดพลาดที่อาจเกิดขึ้นในการปฏิบัติการวัดประเมินพื้นที่ใบเป็นสิ่งที่เกิดขึ้นได้จากวิธีการประเมินพื้นที่ใบด้วยวิธีการข้างต้น การนำเทคโนโลยีทางคอมพิวเตอร์มาประยุกต์สำหรับงานดังกล่าวจึงเป็นสิ่งที่น่าสนใจ เทคนิคการประมวลผลภาพเป็นอีกวิธีการหนึ่งที่ได้รับคามนิยมในการนำมาใช้สร้างระบบและเครื่องมือที่ใช้คอมพิวเตอร์ในการวิเคราะห์ผลลัพธ์ ซึ่งผู้วิจัยได้เลือกใช้วิธีการกำเนิดภาพด้วยการใช้เครื่องกวาดตรวจ(Scanner) และใช้ระบบภาพ RGB [3] ในการดำเนินการ โดยกำหนดค่าความละเอียดที่ใช้ในสแกนภาพที่ความละเอียด200 และ 100 จากนั้นภาพที่ได้จากการสแกนด้วยเครื่องกวาดตรวจจะถูกนำมาวิเคราะห์เพื่อจำแนกส่วนของใบและพื้นที่หลังออกจากกันด้วยอัลกอริทึมที่ออกแบบไว้ ค่าจำนวนของพิกเซลจะถูก

นำมาใช้ในการประเมินขนาดพื้นที่ใบพืชด้วยแบบจำลองที่เหมาะสม โดยจะนำเสนอในหัวข้อต่อจากนี้ [3]

2. การสร้างแบบจำลองในการประเมินพื้นที่ใบ

ในกระบวนการนี้จะนำเสนอขั้นตอนการสร้างแบบจำลองเพื่อใช้สำหรับประเมินขนาดพื้นที่ใบพืชจากจำนวนจุดสีของใบพืชจากภาพ ข้อมูลภาพตัวอย่างที่สร้างจากกระดาษมาตรฐานที่มีขนาดที่แตกต่างกันจำนวน 100 ตัวอย่างถูกนำมาใช้เป็นชุดข้อมูลสำหรับสร้างแบบจำลองเพื่อใช้ในการประเมินขนาดของพื้นที่ใบพืช[4] เพื่อใช้สำหรับหาผลลัพธ์ของจำนวนจุดสีของกระดาษมาตรฐานในแต่ละขนาด โดยข้อมูลจุดสีจากภาพตัวอย่างทั้งหมดแยกตามค่าความละเอียดจะถูกใช้ในการสร้างแบบจำลองสำหรับประเมินขนาดพื้นที่ใบ และวิเคราะห์ความถดถอยเชิงเส้นเป็นวิธีที่ถูกเลือกในการดำเนินการสร้างแบบจำลองครั้งนี้ การวิเคราะห์ความถดถอยเชิงเส้นอย่างง่ายคือ การวิเคราะห์ความสัมพันธ์ของ 2 ตัวแปรคือตัวแปรอิสระหนึ่งตัวและตัวแปรตามหนึ่งตัว ความสัมพันธ์ของตัวแปรทั้งสองเป็นลักษณะเชิงเส้นตรง ถ้ามีตัวแปร 2 ตัวเป็นคู่ๆ เราอาจจะคิดว่าตัวแปร 2 ตัวนั้นมีความสัมพันธ์กันในลักษณะใด โดยการสร้างแผนภาพการกระจาย (Scatter Diagram) ระหว่างตัวแปร 2 ตัว ในกรณีที่แผนภาพการกระจายนั้นเป็นเส้นตรง เราสามารถสร้างเป็นความสัมพันธ์ในรูปแบบของสมการได้ ดังสมการที่ 1 ซึ่งสมการดังกล่าวเรียกว่าสมการถดถอยอย่างง่าย [5]

$$Y_i = \alpha + \beta x_i + \varepsilon_i \quad (1)$$

โดย Y_i เป็นค่าของตัวแปรตาม

α เป็นค่าของ Y_i เมื่อ X_i เป็นศูนย์

β เป็นค่า Slope ของสมการ, สัมประสิทธิ์ของความถดถอย

ε_i เป็นค่าความคลาดเคลื่อนระหว่างค่าจริงกับค่าประมาณของ Y_i

วิธีหาค่า α และค่า β จะประมาณโดยวิธีกำลังสองน้อยที่สุดซึ่งเป็นวิธีที่ได้รับความนิยมกันมาก โดยมีหลักการคือจะทำให้

ผลรวมกำลังสองของความแตกต่างระหว่าง Y_i และ Y ที่ได้จากการประมาณมีค่าต่ำที่สุด ซึ่งสมการประมาณการจะมีรูปแบบดังสมการที่ 2 [4]

$$\hat{Y}_i = a + bx_i \quad (2)$$

โดย a เป็นค่า Y-intercept เป็นค่าประมาณของ α

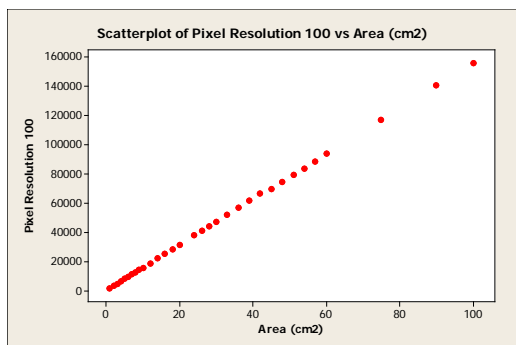
b คือค่า Slope เป็นค่าประมาณของ β

\hat{Y}_i คือค่าประมาณการของ Y_i เมื่อตัวแปรอิสระคือ x_i

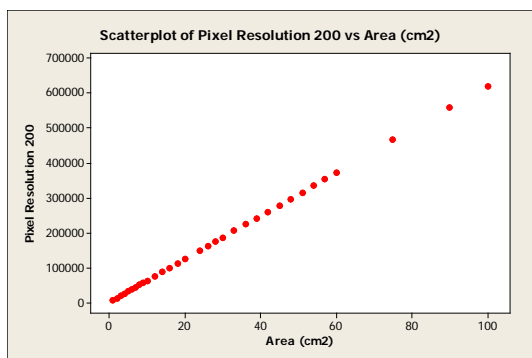
ในการสร้างแบบจำลองด้วยวิธีวิเคราะห์ความถดถอยเชิงเส้นแบบง่ายนั้นผู้วิจัยได้นำเสนอแบบจำลองจำนวน 3 รูปแบบคือแบบจำลองที่สร้างจากจำนวนจุดสีที่ความละเอียด 100 แบบจำลองที่สร้างจากจำนวนจุดสีที่ความละเอียด 200 แบบจำลองที่ใช้ค่าอัตราส่วนระหว่างจำนวนจุดสีของพื้นหลังและพื้นผิวใบ โดยแสดงตัวอย่างข้อมูลในตารางที่ 1 แสดงกราฟแบบกระจายดังภาพที่ 1 และ 2

ตารางที่ 1: แสดงตัวอย่างข้อมูลที่ใช้สร้างแบบจำลอง

Size of the object	Pixel of Leaf at Resolution 100	Pixel of Leaf at Resolution 200
1	1824	6834
2	3533	13315
3	4943	19105
4	6600	25735
5	8261	32055
6	9759	37982
7	11446	44461
8	12724	50671
9	14235	56840
...
75	117019	466890
90	140382	560109
100	155631	620944



ภาพที่ 1: แสดงกราฟแบบกระจายจากการสแกนที่ 100



ภาพที่ 2: แสดงกราฟแบบกระจายจากการสแกนที่ 200

จากข้อมูลเมื่อทำการวิเคราะห์ด้วยกรรมวิธีความถดถอยเชิงเส้นพบว่าจำนวนจุดสีของใบพืชที่การสแกนความละเอียดต่าง ๆ มีความเป็นเส้นตรงแบบแปรผันตรง ซึ่งได้แบบจำลองสำหรับประเมินพื้นที่ใบพืชด้วยการสแกนที่ความละเอียด 100 ดังสมการที่ 3 และได้แบบจำลองสำหรับประเมินพื้นที่ใบพืชด้วยการสแกนที่ความละเอียด 200 ดังสมการที่ 4

$$LeafArea = -.299 + (.000644 \times P) \tag{3}$$

$$LeafArea = -.111 + (.000161 \times P) \tag{4}$$

โดย P คือจำนวนจุดสีที่ความละเอียด 200 และ 100

จากแบบจำลองสำหรับประเมินพื้นที่ใบพืชด้วยการสแกนที่ความละเอียด 100 และ 200 จะเห็นได้ว่า หากทำการสแกนที่ความละเอียดต่างกัน จะส่งผลต่อแบบจำลองที่เปลี่ยนแปลงไป หรือกล่าวได้ว่าแบบจำลองจะยึดโยงกับความละเอียดในการสแกน โดยไม่มีความยึดหยุ่นกับค่าความละเอียดอื่น ๆ แบบจำลองที่ใช้ค่าอัตราส่วนระหว่างจำนวนจุดสีของพื้นหลังและพื้นผิวใบ จึงถูกสร้างเพื่อแก้ไขข้อจำกัด

ดังกล่าวเพื่อให้มีคุณสมบัติที่ปรับเหมาะตามค่าความละเอียดที่เปลี่ยนแปลง โดยแสดงตัวอย่างข้อมูลในการสร้างแบบจำลองในตารางที่ 2 และแสดงแบบจำลองดังสมการที่ 5

ตารางที่ 2: แสดงตัวอย่างข้อมูลที่ใช้สร้างแบบจำลองจากค่าอัตราส่วน

object's actual size (cm ²)	Pixel of Leaf Resolution	Pixel of Background Resolution	Ratio of Leaf (R)
1	6834	4084150	0.17
2	13315	4077669	0.33
3	19105	4071879	0.47
4	25735	4065249	0.63
...
100	620944	3470040	15.18

$$Leafarea = -.110 + (6.60 \times R) \tag{5}$$

โดย R คือค่าอัตราส่วนระหว่างจุดสีพื้นหลังและพื้นผิวใบ

3. ทดสอบและหาความคลาดเคลื่อนของแบบจำลอง

จากแบบจำลองทั้งสามที่นำเสนอผู้วิจัยได้นำแบบจำลองมาทำการทดสอบการประเมินพื้นที่ใบเพื่อหาค่าความคลาดเคลื่อนของแบบจำลองทั้งสาม โดยใบพืชจะถูกตรวจวัดประเมินขนาดพื้นที่ใบด้วยเครื่องวัดพื้นที่ใบ (WinDIAS Colour Image) ตัวอย่างใบพืชที่มีลักษณะแตกต่างกันจำนวน 60 ใบ จะถูกนำมาใช้ในการหาค่าความคลาดเคลื่อนในการประเมินพื้นที่ใบ โดยแสดงผลการประเมินด้วยแบบจำลองทั้งสามในตารางที่ 3

ตารางที่ 3: แสดงผลการทดสอบแบบจำลองทั้งสาม 3

No.	Model Resolution 100	Model Resolution 200	Model Ratio of Leaf	Actual size of leaf
1	0.876	0.989	0.992	1
2	1.976	2.033	2.038	2

ตารางที่ 3: แสดงผลการทดสอบแบบจำลองทั้ง 3 (ต่อ)

No.	Model Resolution 100	Model Resolution 200	Model Ratio of Leaf	Actual size of leaf
3	3.951	4.032	4.041	4
4	13.977	13.935	13.964	14
...
60	99.927	99.861	100.067	100

จากตารางที่ 3 เป็นผลการทดสอบแบบจำลองทั้งสาม 3 โดยผลการทดลองพบว่าแบบจำลองสำหรับประเมินพื้นที่ใบพืชด้วยค่าอัตราส่วนจุดสี่มีความคลาดเคลื่อนต่ำที่สุด โดยมีค่าความคลาดเคลื่อนเฉลี่ยเท่ากับ 437.% และแบบจำลองสำหรับประเมินพื้นที่ใบพืชด้วยการสแกนที่ความละเอียด 200 มีค่าความคลาดเคลื่อนเฉลี่ยเท่ากับ 457.% และแบบจำลองสำหรับประเมินพื้นที่ใบพืชด้วยวิธีการสแกนที่ความละเอียด 100 มีค่าความคลาดเคลื่อนเฉลี่ยเท่ากับ 1.098% สรุปผลการทดสอบความคลาดเคลื่อนเฉลี่ยของทุกแบบจำลองที่นำเสนอในตารางที่ 4

ตารางที่ 4: แสดงสรุปผลการทดสอบความคลาดเคลื่อนเฉลี่ย

No.	Type	Model	Error Average
1	Resolution 100	$Leaf_{area} = -.299 + (.000644 \times P)$	1.098%
2	Resolution 200	$Leaf_{area} = -.111 + (.000161 \times P)$.457%
3	Ratio of Leaf	$Leaf_{area} = -.110 + (6.60 \times R)$.437%

4. สรุปผล

จากแบบจำลองทั้งสามที่นำเสนอพบว่าแบบจำลองที่มีค่าความคลาดเคลื่อนที่ต่ำที่สุดคือแบบจำลองสำหรับประเมินพื้นที่ใบพืชด้วยค่าอัตราส่วนระหว่างจุดสี่พื้นหลังและพื้นผิวใบ นอกจากนี้ผู้วิจัยพบว่าแบบจำลองดังกล่าวมีความยืดหยุ่นต่อการเปลี่ยนแปลงค่าความละเอียด และพบว่าการสแกนที่ค่าความละเอียดที่สูงขึ้นจะส่งผลต่อความคลาดเคลื่อนที่ลดลง

เอกสารอ้างอิง

- [1] Wongchun W. Principles of plant physiology. Bangkok, Funny publishing; 1998
- [2] Parparshut P. Physiology of plants. Bangkok. Odeon store. Bangkok;2007
- [3] Karpech P. Finding the leaf areas from digital photos. Khon Kaen Agriculture Journal; 2011,p.392 – 397
- [4] Thaiparnit S. Design algorithms and modeling to determine the leaf area from photos. 2nd International Conference on Education Reform and Management Innovation; 2013.p.
- [5] บุญใจ ศรีสถิตินรากร .ระเบียบวิธีวิจัยแนวทางสู่ความสำเร็จ . สำนักพิมพ์แห่งจุฬาลงกรณ์กรุงเทพฯ .;2550

ซัพพอร์ตเวกเตอร์แมชชีนสำหรับการพยากรณ์ราคาตราสารอนุพันธ์

Support Vector Machines for Derivatives Price Prediction

สุวัชร ภิญโญพันธ์ (Suwat Pinyopan)¹ และ บุญเสริม กิจศิริกุล (Boonserm Kijirikul)²

ภาควิชาวิศวกรรมคอมพิวเตอร์ จุฬาลงกรณ์มหาวิทยาลัย

¹suwat1982@gmail.com, ²boonserm.k@chula.ac.th

บทคัดย่อ

บทความนี้นำเสนอวิธีการสำหรับพยากรณ์ราคาตราสารอนุพันธ์ โดยใช้ซัพพอร์ตเวกเตอร์แมชชีนในการพยากรณ์และใช้RReliefFสำหรับคัดเลือกคุณลักษณะที่เหมาะสม วิธีนี้เริ่มต้นโดยการรวบรวมคุณลักษณะจากข้อมูลที่เกี่ยวข้องกับการพยากรณ์ราคาตราสารอนุพันธ์แล้วใช้ RReliefFกรองเอาเฉพาะคุณลักษณะที่สำคัญๆ จากนั้นจึงสร้างข้อมูลฝึกและทดสอบจากคุณลักษณะที่คัดกรองได้ แล้วฝึกโมเดลของซัพพอร์ตเวกเตอร์แมชชีนและวัดผล ผลการทดลองที่ได้แสดงให้เห็นว่าวิธีการที่นำเสนอให้ผลตอบแทนการลงทุนที่ดีกว่าวิธีการดั้งเดิมแบบ Buy-and-Hold

คำสำคัญ: ซัพพอร์ตเวกเตอร์แมชชีนการพยากรณ์ราคาตราสารอนุพันธ์

Abstract

This paper presents a method for derivative price prediction. The method is based on Support Vector Machines, and employs RReliefF for attribute selection. First, several attributes are collected from the available data relevant to derivative price prediction. The most important attributes are then selected from these attributes by RReliefF. Then training and test examples are constructed based on the selected attributes. An SVM model is then learned and its prediction results are evaluated. The results show that the proposed method yields higher investment return than the traditional Buy-and-Hold strategy.

Keywords: Support Vector Machines, Derivatives Price Prediction.

1. บทนำ

การซื้อขายตราสารทางการเงินให้ได้ผลตอบแทนตามที่ผู้ลงทุนต้องการนั้น ต้องอาศัยการวิเคราะห์แนวโน้มการเคลื่อนไหวของราคาที่แม่นยำ และช่วงเวลาในการตัดสินใจซื้อและขายที่เหมาะสม การวิเคราะห์ตราสารทางการเงินที่ใช้ส่วนมากมีสองวิธีคือ 1.การวิเคราะห์ปัจจัยพื้นฐาน คือ การวิเคราะห์จากอัตราส่วนทางการเงิน ประกอบด้วยปัจจัยภายนอก เช่น สภาพเศรษฐกิจ และการเมืองของประเทศ 2. การวิเคราะห์เชิงเทคนิค คือ การวิเคราะห์จากข้อมูลการซื้อขายในช่วงเวลาที่ผ่านมา และอาศัยตัวบ่งชี้(indicator) เป็นเครื่องมือช่วยในการตัดสินใจซื้อขาย นอกจากการวิเคราะห์ทั้ง 2 วิธีแล้ว มีผู้ศึกษาการใช้ปัญญาประดิษฐ์(Artificial Intelligence)มาประยุกต์ใช้กับการพยากรณ์ราคาหรือการซื้อขายตราสารทางการเงิน เช่น ข่ายงานประสาทเทียม (Artificial Neuron Network)[1] ขั้นตอนวิธีเชิงพันธุกรรม(Genetic Algorithm)[2]เป็นต้น งานวิจัยที่เกี่ยวข้องกับการพยากรณ์ตราสารทางการเงินส่วนมากเป็นการพยากรณ์ราคาในวันถัดไป แต่ในงานวิจัยนี้จะทำการพยากรณ์ราคาตราสารทางการเงินในอนาคตที่ถัดไปซึ่งเป็นแนวคิดใหม่ในการลงทุนที่เรียกว่าการซื้อขายความถี่สูง (High Frequency Trading)และทำการทดสอบผลการพยากรณ์ด้วยการจำลองการซื้อขาย เพื่อเป็นแนวทางในการพัฒนาระบบการซื้อขายตราสารทางการเงินในระยะสั้นต่อไป

2. วรรณกรรมที่เกี่ยวข้อง

มีงานวิจัยที่เกี่ยวข้องมากมายที่ศึกษาเรื่องการประยุกต์ใช้ปัญญาประดิษฐ์ในการซื้อขายตราสารทางการเงินด้วยระบบคอมพิวเตอร์Kablan[3]ศึกษาการใช้ Adaptive Neuro Fuzzy Inference Systemในพยากรณ์และซื้อขายเงินตราต่างประเทศ

ข้อมูลที่ใช้เป็นอัตราแลกเปลี่ยนเงินตราต่างประเทศ EUR/USD ที่มีการเปลี่ยนแปลงทุกๆ 5 นาที ผลการทดลองพบว่าเมื่อเปรียบเทียบกับวิธี Buy-and-Hold[4]แล้วได้ผลตอบแทนมากกว่า Bao และ คณะ [5] ทำการศึกษาการพยากรณ์ดัชนีตลาดหลักทรัพย์ในวันถัดไปด้วย Fuzzy Support Vector Machines Regression ผลการทดลองเมื่อเปรียบเทียบกับ การพยากรณ์ด้วยซัพพอร์ตเวกเตอร์แมชชีนแบบมาตรฐานมีค่า Normalized Mean Square Error น้อยกว่า อย่างไรก็ตาม วิจัยไม่ได้วัดผลการทำนายด้วยการซื้อขาย พิตติพล คันธวัฒน์ [6] ศึกษาเรื่องการผสมการวิเคราะห์เชิงเทคนิคและซัพพอร์ตเวกเตอร์แมชชีนสำหรับการซื้อขายหลักทรัพย์โดยทำการเปลี่ยนแปลงราคาหุ้นในวันถัดไป ซึ่งการเปลี่ยนแปลงมี 3 ลักษณะคือ 1.เพิ่มขึ้น 2.คงที่ 3.ลดลง และจำลองการซื้อขายหุ้นตามผลการทำนายราคาในวันถัดไป ผลการทดลองพบว่าได้ผลกำไรจากการซื้อขายหุ้นมากกว่าวิธีการดั้งเดิม

ในงานวิจัยนี้เราต้องการพยากรณ์ราคาตราสารอนุพันธ์ใน นาทีถัดไปโดยการใช้คุณสมบัติที่เกี่ยวข้องจำนวน 18 ตัว และทดสอบประสิทธิภาพในการทำนายราคาด้วยการซื้อขายจำลอง ดังจะกล่าวรายละเอียดดังต่อไปนี้

3. การพยากรณ์ราคาตราสารอนุพันธ์ด้วยซัพพอร์ตเวกเตอร์แมชชีน

3.1 ซัพพอร์ตเวกเตอร์แมชชีน

ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machines: SVM) ได้รับการพัฒนาจาก Vapnik [7] สามารถใช้ได้กับการจำแนกประเภท (Classify) และการถดถอย (Regression) มีการนำไปใช้อย่างแพร่หลาย รวมถึงการพยากรณ์ราคาตราสารทางการเงิน เพราะสามารถทำงานได้เช่นเดียวกับเครือข่ายประสาทเทียมและขั้นตอนวิธีเชิงพันธุกรรม แต่ซัพพอร์ตเวกเตอร์แมชชีนมีพารามิเตอร์ที่ต้องปรับน้อยกว่า คือ มีเพียงฟังก์ชันเคอร์เนลที่ใช้ในการฝึกและค่าสัมประสิทธิ์บางตัวเท่านั้น ในขณะที่เครือข่ายประสาทเทียมต้องเลือกชนิดที่ใช้ จำนวนชั้นซ่อน จำนวนปมในแต่ละชั้น กฎการเรียนรู้ อัตราการเรียนรู้ โมเมนตัม ซึ่งมีความยุ่งยากนำไปใช้มากกว่า และเมื่อเปรียบเทียบกับขั้นตอนวิธีเชิงพันธุกรรมแล้ว ผลการพยากรณ์ของซัพพอร์ตเวกเตอร์แมชชีนในแต่ละครั้งด้วยข้อมูลทดสอบชุดเดียวกันจะ

ได้ผลลัพธ์เหมือนกันทุกครั้ง แต่เมื่อใช้ขั้นตอนวิธีเชิงพันธุกรรมด้วยข้อมูลชุดเดียวกันจะได้ผลลัพธ์ในแต่ละครั้งแตกต่างกัน

3.2 การถดถอยด้วยซัพพอร์ตเวกเตอร์แมชชีน (SVM Regression - SVMR)

ซัพพอร์ตเวกเตอร์แมชชีนแบบจำแนกประเภทข้อมูลนั้นสามารถแบ่งข้อมูลเป็นประเภท+หรือ- ได้ โดยใช้ +1 แทนข้อมูลประเภท+ และใช้ตัวเลข -1 แทนข้อมูลประเภท- ส่วนงานวิจัยนี้เราต้องการนำซัพพอร์ตเวกเตอร์แมชชีนมาใช้ในการทำนายค่าตัวเลขที่เป็นจำนวนจริง ซึ่งทำได้ด้วยการใช้ฟังก์ชันสูญเสียแบบ ϵ -insensitive (ϵ -insensitive loss function) เข้ามาช่วยในการสร้าง $f(x_k)$ ที่ทุกๆ ข้อมูลการฝึกมีความเบี่ยงเบนจากค่า y ที่เป็นเป้าหมายตามเงื่อนไข $(w \cdot x_i + b) - y_i \leq \epsilon$ และ $y_i - (w \cdot x_i + b) \leq \epsilon$ สำหรับ $i = 1, \dots, l$ ซัพพอร์ตเวกเตอร์แมชชีนแบบนี้เรียกว่าการถดถอยด้วยซัพพอร์ตเวกเตอร์แมชชีนหรือ SVMR แบบระยะขอบแข็ง (hard margin) เราต้องทำให้ค่า $\frac{1}{2} \|w\|^2$ น้อยที่สุดและ SVMR แบบระยะขอบอ่อน (soft margin) โดยต้องทำให้ค่า $\frac{1}{2} \|w\|^2 - C \sum_{i=1}^l (\xi_i^+ + \xi_i^-)$ ให้มีค่าน้อยที่สุด โดยมีเงื่อนไขเพิ่มเติมคือ

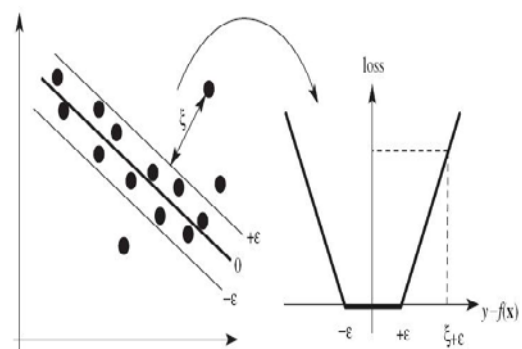
$$(w \cdot x_i + b) - y_i \leq \epsilon + \xi_i^+$$

$$y_i - (w \cdot x_i + b) \leq \epsilon + \xi_i^-$$

$$\xi_i^+ \geq 0 \text{ และ } \xi_i^- \geq 0 \text{ สำหรับ } i = 1, \dots, l \quad (1)$$

ξ_i เป็นตัวปรับข้อมูลที่ผิดพลาด เรียกว่า ตัวแปรหย่อน (Slack variable)

การทำงานของ SVMR แสดงดังภาพที่ 1



ภาพที่ 1: SVMR แบบระยะขอบอ่อน และ ϵ -insensitive loss function

จะได้สมการลากรอง โดย $\alpha_i^+, \alpha_i^-, \mu_i^-, \mu_i^-$ คือตัวคูณลากรอง

$$L = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l (\xi_i^+ + \xi_i^-) - \sum_{i=1}^l (\mu_i^+ \xi_i^+ + \mu_i^- \xi_i^-) - \sum_{i=1}^l \alpha_i^+ (\varepsilon + \xi_i^+ + y_i - w \cdot x_i - b) - \sum_{i=1}^l \alpha_i^- (\varepsilon + \xi_i^- - y_i + w \cdot x_i + b) \quad (2)$$

จากนั้นทำการหาค่าต่ำสุดเมื่อเทียบกับ w, b, ξ_i^+, ξ_i^- บนเงื่อนไข $\xi_i^+, \xi_i^-, \alpha_i^+, \alpha_i^-, \mu_i^+, \mu_i^- \geq 0$

$$\frac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^l (\alpha_i^+ - \alpha_i^-) = 0 \quad (3)$$

$$\frac{\partial L}{\partial w} = 0 \rightarrow w = \sum_{i=1}^l (\alpha_i^- - \alpha_i^+) x_i \quad (4)$$

$$\frac{\partial L}{\partial \xi_i^+} = 0 \rightarrow C - \alpha_i^+ - \mu_i^+ = 0 \quad (5)$$

$$\frac{\partial L}{\partial \xi_i^-} = 0 \rightarrow C - \alpha_i^- - \mu_i^- = 0 \quad (6)$$

จะได้สมการลากรองคือ

$$L = -\frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l (\alpha_i^- - \alpha_i^+) (\alpha_j^- - \alpha_j^+) (x_i \cdot x_j) \quad (7)$$

ปัญหาการหาค่าเหมาะสมที่สุดจะเป็นการทำให้ $-\varepsilon \sum_{i=1}^l (\alpha_i^- - \alpha_i^+) + \sum_{i=1}^l y_i (\alpha_i^- - \alpha_i^+)$ ให้มากที่สุดตามเงื่อนไข $\sum_{i=1}^l (\alpha_i^- - \alpha_i^+) = 0$ และ $\alpha_i^-, \alpha_i^+ \in [0, C]$ และ $i = 1, \dots, l$ ซึ่งจะนำไปสู่ฟังก์ชันค่า

$$f(x_k) = \sum_{i=1}^l (\alpha_i^- - \alpha_i^+) x_i \cdot x_k + b \quad (8)$$

3.3 ขั้นตอนวิธีRRReliefF

ขั้นตอนวิธีRRReliefF[8]เป็นวิธีการประเมินคุณภาพของคุณลักษณะที่ใช้ในการฝึก เพื่อเพิ่มประสิทธิภาพในการฝึกของการเรียนรู้ของเครื่อง เนื่องจากคุณลักษณะหากมีมากเกินไปอาจส่งผลให้การฝึกมีประสิทธิภาพลดลง ดังนั้นจึงจำเป็นต้องเลือกเฉพาะคุณลักษณะเท่าที่จำเป็นต่อการเรียนรู้ปัญหา

วิธีการนี้จะสร้าง $W[A]$ สำหรับประเมินคุณภาพของคุณลักษณะ A โดยการประมาณความแตกต่างของค่าความน่าจะเป็น ดังนี้

$$W[A] = P(\text{different value of } A \mid \text{nearest instant from different class}) - P(\text{different value of } A \mid \text{nearest instant from same class}) \quad (9)$$

ซึ่งก็คือการให้คะแนนน้ำหนักเพิ่มหากคุณลักษณะนั้นมีส่วนในการแบ่งแยกประเภทของตัวอย่างที่กำลังพิจารณาอยู่และลดคะแนนน้ำหนักหากคุณลักษณะนั้นไม่มีส่วนในการแบ่งแยกประเภทของตัวอย่างนั้นรหัสเทียบอย่างคร่าวๆ ของขั้นตอนวิธีRRReliefF เป็นดังภาพที่2ด้านล่างนี้

```

1.กำหนดค่าเริ่มต้นของNdC,NdA[A],
   NdC&dA[A],W[A]ให้เป็น 0;
2. Fori := 1 to m Do
   2.1 สุ่มเลือกตัวอย่างฝึกRi;
   2.2 เลือกตัวอย่างฝึกจำนวนkตัวIj ที่ใกล้กับRi;
   2.3 Forj := 1 to k Do
     2.3.1 NdC := NdC + diff(T(.), Ri, Ij) . d(i, j);
     2.3.2 ForA := 1 to a Do
       2.3.2.1 NdA[A] := NdA[A] + diff(A, Ri, Ij) × d(i, j);
       2.3.2.2 NdC&dA[A] := NdC&dA[A] + diff(T(.),
         Ri, Ij) × diff(A, Ri, Ij) × d(I, j);

```

ภาพที่2: ขั้นตอนวิธีRRReliefF

ในภาพที่ 2 นั้น $T(x)$ เป็นค่าคำตอบที่ต้องการทำนาย และพจน์ $d(i, j)$ เป็นการวัดระยะห่างระหว่างตัวอย่าง R_i และ I_j ซึ่งนิยามดังนี้

$$d(i, j) = \frac{d_1(i, j)}{\sum_{l=1}^k d_1(i, l)}$$

และ

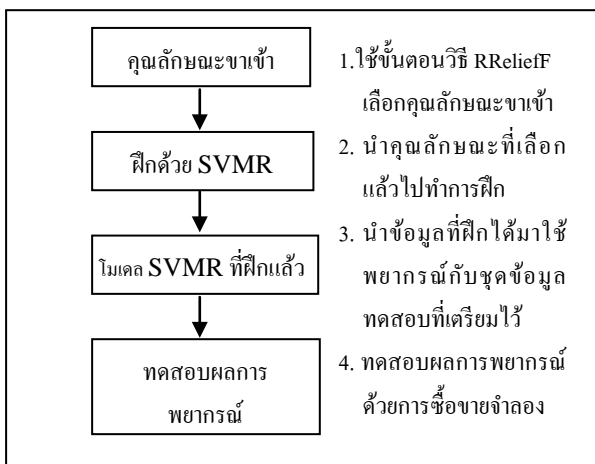
$$d_1(i, j) = \exp\left(-\frac{\text{rank}(R_i, I_j)}{\sigma}\right)^2 \quad (10)$$

$\text{rank}(R_i, I_j)$ เป็นอันดับของระยะห่างเรียงจากน้อยไปหามาก และ σ เป็นตัวแปรที่กำหนดโดยผู้ใช้งานที่ต้องการให้ความห่างมีผลมากน้อยเพียงใด ฟังก์ชัน $\text{diff}(A, R_i, I_j)$ คำนวณค่าความต่างระหว่างค่าของคุณลักษณะ A ของตัวอย่าง R_i และ I_j ดังด้านล่างนี้

$$\text{diff}(A, R_i, I_j) = \frac{|\text{value}(A, R_i) - \text{value}(A, I_j)|}{\max(A) - \min(A)} \quad (11)$$

3.4 การพยากรณ์ราคาตราสารอนุพันธ์ด้วย SVMR

ขั้นตอนการพยากรณ์ราคาตราสารอนุพันธ์ด้วย SVMR ในงานวิจัยนี้ แสดงในภาพที่ 3



ภาพที่ 3: ขั้นตอนการฝึกด้วย SVMR

เริ่มจากการใช้ขั้นตอนวิธี RReliefF เพื่อคัดเลือกคุณลักษณะขาเข้าที่เหมาะสมในการฝึก จากนั้นนำข้อมูลที่ได้ออกไปฝึก SVMR ผลที่ได้จึงนำไปใช้ในการทำนายข้อมูลในอนาคตหรือข้อมูลทดสอบ งานวิจัยนี้ได้ทำการซื้อขายจำลองและวัดผลการพยากรณ์ โดยมีรายละเอียดดังต่อไปนี้

3.5 ขั้นตอนการทดลอง

การเตรียมคุณลักษณะขาเข้าสำหรับฝึก

บทความนี้เลือกใช้คุณลักษณะขาเข้าที่คาดว่าจะช่วยในการพยากรณ์ราคาตราสารอนุพันธ์มาทั้งสิ้น 18 ตัวซึ่งโดยทั่วไปในงานวิจัยอื่นมักใช้ ราคาปัจจุบัน ราคาต่ำสุด ราคาสูงสุด ราคาย้อนหลัง เป็นหลัก แต่เราได้เพิ่มค่าเฉลี่ยเคลื่อนที่ และดัชนีกำลังสัมพัทธ์ย้อนหลัง มาช่วยด้วย

อย่างไรก็ดีเรายังได้นำเสนอการใช้ RReliefF มาคัดเลือกเฉพาะคุณลักษณะที่สำคัญจริงๆ จากคุณลักษณะทั้ง 18 ตัวนี้จึงจะแสดงผลที่ได้ต่อไปคุณลักษณะขาเข้าที่ใช้ในการฝึกแสดงดังตารางที่ 1

ตารางที่ 1: คุณลักษณะที่ใช้ในการฝึก

	คุณลักษณะขาเข้า
X_1	ราคานาทีปัจจุบัน
X_2	ค่าเฉลี่ยเคลื่อนที่ของราคา 5 นาที
X_3	ค่าเฉลี่ยเคลื่อนที่ของราคา 10 นาที
X_4	ค่าเฉลี่ยเคลื่อนที่ของราคา 15 นาที
X_5	ค่าเฉลี่ยเคลื่อนที่ของราคา 20 นาที
X_6	ค่าเฉลี่ยเคลื่อนที่ของราคา 25 นาที
X_7	ค่าเฉลี่ยเคลื่อนที่ของราคา 30 นาที
X_8	(ดัชนีกำลังสัมพัทธ์ย้อนหลัง 5 นาที - 100) / 100
X_9	(ดัชนีกำลังสัมพัทธ์ย้อนหลัง 10 นาที - 100) / 100
X_{10}	(ดัชนีกำลังสัมพัทธ์ย้อนหลัง 15 นาที - 100) / 100
X_{11}	(ดัชนีกำลังสัมพัทธ์ย้อนหลัง 20 นาที - 100) / 100
X_{12}	ราคา 5 นาทีก่อนหน้า
X_{13}	ราคา 10 นาทีก่อนหน้า
X_{14}	ราคา 15 นาทีก่อนหน้า
X_{15}	ราคา 20 นาทีก่อนหน้า
X_{16}	ราคาเปิดนาทีปัจจุบัน
X_{17}	ราคาต่ำสุดในนาทีปัจจุบัน
X_{18}	ราคาสูงสุดในนาทีปัจจุบัน
	ข้อมูลขาออก
Y	ราคาในอนาคตถัดไป

ค่าเฉลี่ยเคลื่อนที่ (Moving Average: MA) คำนวณจากสูตรดังนี้ $MA = \frac{a_1 + a_2 + a_3 + \dots + a_n}{n}$ (11)

a_1 คือ ราคาตราสารอนุพันธ์ก่อนหน้า 1 นาทีที่ผ่านมา a_2 คือ ราคาตราสารอนุพันธ์ก่อนหน้า 2 นาทีที่ผ่านมา a_3 คือ ราคาตราสารอนุพันธ์ก่อนหน้า 3 นาทีที่ผ่านมา a_n คือ ราคาตราสารอนุพันธ์ก่อนหน้า n นาทีที่ผ่านมา n คือ จำนวนนาที

ดัชนีกำลังสัมพัทธ์ย้อนหลัง (Relative Strength Index : RSI) คำนวณจากสูตร

$$RSI = 100 \left(\frac{U}{U + D} \right) \quad (12)$$

โดยที่ U เป็นค่าเฉลี่ยของจำนวนที่เปลี่ยนแปลงเพิ่มขึ้นของราคาปิดใน 14 นาที, D เป็นค่าเฉลี่ยของจำนวนที่เปลี่ยนแปลงลดลงของราคาปิดใน 14 นาที

เมื่อทำการคัดเลือกคุณลักษณะด้วยขั้นตอนวิธีRReliefF แล้ว จะเหลือคุณลักษณะที่จำเป็นเพียงแค่ 8คุณลักษณะเท่านั้น ดังตารางที่ 2

ตารางที่ 2:คุณลักษณะที่ผ่านการคัดเลือกโดย RReliefF

คุณลักษณะขาเข้า	
X_1	ราคานาทีปัจจุบัน
X_2	(ดัชนีกำลังสัมพัทธ์ย้อนหลัง 5นาที- 100) / 100
X_3	(ดัชนีกำลังสัมพัทธ์ย้อนหลัง 10นาที- 100) / 100
X_4	(ดัชนีกำลังสัมพัทธ์ย้อนหลัง 15นาที- 100) / 100
X_5	(ดัชนีกำลังสัมพัทธ์ย้อนหลัง 20นาที- 100) / 100
X_6	ราคาเปิดนาทีปัจจุบัน
X_7	ราคาต่ำสุดในนาทีปัจจุบัน
X_8	ราคาสูงสุดในนาทีปัจจุบัน
ข้อมูลขาออก	
Y	ราคาในนาทีถัดไป

เมื่อได้คุณลักษณะที่คัดเลือกแล้ว จึงนำไปฝึกด้วยซอฟต์แวร์ mySVM [9]โดยใช้ฟังก์ชันเคอร์เนลคือRBFด้วย default parameters

เมื่อทำการฝึกเสร็จแล้วจึงนำโมเดลที่ฝึกได้มาทำการพยากรณ์ด้วยข้อมูลที่เตรียมไว้ จากนั้นนำผลการพยากรณ์มาทำการซื้อขายจำลอง โดยมีขั้นตอนในการซื้อขายคือ เมื่อราคาที่ทำพยากรณ์ได้ในนาทีถัดไปตกลงจะทำการซื้อ และถ้าราคาที่ทำพยากรณ์ได้ในนาทีถัดไปเพิ่มขึ้นหนึ่งจุดจะทำการขาย

3.6การประเมินผล

การประเมินผลการทดลองของงานวิจัยนี้แบ่งออกเป็น 2 ส่วนคือ 1.วัดค่าคลาดเคลื่อนระหว่างราคาจริงกับราคาที่ทำพยากรณ์ได้ ด้วยค่าคลาดเคลื่อนเฉลี่ยสัมบูรณ์(Mean Absolute Error: MAE) และค่าคลาดเคลื่อนกำลังสอง(Mean Square Error: MSE)2.วัดผลตอบแทนจากการซื้อขายตราสารอนุพันธ์เป็นเวลา 30วันเป็นเปอร์เซ็นต์โดยการนำราคาที่ทำพยากรณ์ได้มาทำการซื้อขายจำลอง ค่าMAEและ MSEแสดงดังสมการ(22)และ (23)ต่อไปนี้

$$MAE = \sum_{i=1}^n \frac{|PV - AV|}{n} \tag{13}$$

$$MSE = \sum_{i=1}^n \frac{(PV - AV)^2}{n} \tag{14}$$

โดยที่PVและAVเป็นราคาที่ทำนายและราคาจริงตามลำดับn เป็นจำนวนข้อมูล

4.ผลการทดลอง

งานวิจัยนี้ใช้ชุดข้อมูลของราคาสัญญาซื้อขายล่วงหน้า SET 50 FUTURES BCMทุกๆ 1นาที่ ระหว่างเดือนกุมภาพันธ์ ถึง เดือนเมษายน2555 โดยแบ่งข้อมูลตั้งแต่ กุมภาพันธ์ 2555 ถึง เดือนมีนาคม 2555 เป็นชุดข้อมูลสำหรับฝึกจำนวน 800ตัวอย่าง และแบ่งข้อมูลสำหรับทดสอบ คือ ข้อมูลการซื้อขายตั้งแต่เดือนมีนาคม 2555 ถึงเดือนเมษายน 2555 (ไม่ซ้ำซ้อนกับชุดข้อมูลฝึก) ซึ่งแบ่งเป็น 3 ชุด คือ 1.ชุดข้อมูลแนวโน้มขาลง(downtrend) จำนวน 201ตัวอย่าง2.ชุดข้อมูลแนวโน้มไร้ทิศทาง(sideways)จำนวน 244ตัวอย่าง และ 3.ชุดข้อมูลแนวโน้มขาขึ้น(uptrend)จำนวน 249 ตัวอย่างผลการทดลองแสดงดังตารางที่ 3

ตารางที่3: เปรียบเทียบค่าคลาดเคลื่อนของข้อมูลทดสอบ

ชุดข้อมูลทดสอบ		ค่าคลาดเคลื่อน	
		MAE	MSE
แนวโน้มขาลง	ไม่ใช้	6.156	40.33
	RReliefF	0.305	0.194
แนวโน้มไร้ทิศทาง	ไม่ใช้	24.106	582.19
	RReliefF	2.144	4.791
แนวโน้มขาขึ้น	ไม่ใช้	3.391	14.013
	RReliefF	0.228	0.119

จากตารางที่ 3 พบว่าค่าคลาดเคลื่อนจากการใช้ RReliefFในการคัดเลือกคุณลักษณะก่อนการฝึกมีค่าน้อยกว่า การฝึกโดยไม่ใช้ RReliefFซึ่งแสดงให้เห็นว่าคุณลักษณะที่ถูกคัดกรองด้วย RReliefFให้ประสิทธิภาพในการเรียนรู้ได้ดีกว่า ดังนั้นจึงนำโมเดลที่เรียนรู้ได้โดยRReliefFมาทำการซื้อขายจำลอง

ตารางที่ 4: เปรียบเทียบกำไรของการลงทุน

ข้อมูลทดสอบ	ผลตอบแทน(จุด)	
	Buy-and-Hold	วิธีที่นำเสนอ
แนวโน้มขาลง	-7.0	-4.0
แนวโน้มไร้ทิศทาง	-1.4	-1.4
แนวโน้มขาขึ้น	13.8	14.1

ผลในตารางที่ 4 แสดงให้เห็นว่าวิธีการที่นำเสนอให้การพยากรณ์ที่ดีกว่าและทำให้ได้กำไรจากซื้อขายในแนวโน้มขาขึ้นด้วยผลตอบแทนที่มากกว่า ในขณะที่แนวโน้มขาลงช่วยให้ขาดทุนลดลง ส่วนแนวโน้มไร้ทิศทางให้ผลตอบแทนไม่ต่างกับวิธี Buy-and-Hold

5.สรุป

งานวิจัยนี้ใช้ซัพพอร์ตเวกเตอร์แมชชีนในการทำนายราคาตราสารอนุพันธ์ในอนาคตโดยมีวัตถุประสงค์เพื่อทดสอบประสิทธิภาพในการพยากรณ์ของซัพพอร์ตเวกเตอร์แมชชีน ผลการทดลองในงานนี้พบว่าการใช้ซัพพอร์ตเวกเตอร์แมชชีนในการพยากรณ์การซื้อขายตราสารอนุพันธ์ทำให้ได้ผลตอบแทนมากกว่าวิธีการซื้อขายแบบ Buy-and-Hold ซึ่งแสดงให้เห็นถึงศักยภาพของวิธีการที่นำเสนอ ซึ่งอาจนำไปปรับปรุงเพื่อใช้ในการซื้อขายความถี่สูงสำหรับตราสารอนุพันธ์ต่างๆ ต่อไปในอนาคต

เอกสารอ้างอิง

- [1] J. S. Zirilli, *Financial Prediction Using Neural Networks*, International Thomson Computer Press, 1996.
- [2] C. Zhou, L. Yu, T. Huang, S. Wang, and K. Lai, "Selecting valuable stock using genetic algorithm", *Lecture Notes in Computer Science*, Vol. 4247, pp. 688-694, 2006.

- [3] A.Kablan, "Adaptive neurofuzzy inference systems for high frequency financial trading and forecasting", *The Third International Conference on Advanced Engineering Computing and Applications in Sciences*, pp. 105-110, 2009.
- [4] http://en.wikipedia.org/wiki/Buy_and_hold
- [5] Y.K. Bao, Z.T. Liu, L. Guo, and W. Wang, "Forecasting stock composite index by fuzzy support vector machines regression", *In Proceedings of 2005 International Conference on Machine Learning and Cybernetics*, pp. 3535-3540, 2005.
- [6] พิตติพล คันธวัฒน์, การผสมการวิเคราะห์เชิงเทคนิคและซัพพอร์ตเวกเตอร์แมชชีนสำหรับการซื้อขายหลักทรัพย์, วิทยานิพนธ์ปริญญาโท, ภาควิชาวิศวกรรมคอมพิวเตอร์, จุฬาลงกรณ์มหาวิทยาลัย, 2551.
- [7] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer-Verlag, New York, 1995.
- [8] R. Sikonja and I. Kononenko, "An adaptation of relief for attribute estimation in regression", *In Proceedings of the Fourteenth International conference on Machine Learning*, Morgan Kaufmann, pp. 296-304, 1997.
- [9] <http://www.ai.cs.unidortmund.de/SOFTWARE/MYSVM/index.html>

พยากรณ์ราคาหลักทรัพย์ด้วยกราฟแท่งเทียนและฮิดเดนมาร์คอฟโมเดล

Stock Forecasting with Candlestick Chart and Hidden Markov Models

สุภกิจ นุตยะสกุล (Supakit Nootyaskool)¹ มณสินี อรรถศิริปัญญา (Monsinee Attasiripanya)²

และ วิสินี ประสิทธิ์ชัยกุล (Wisinee Prasaditichaikul)³

คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

¹Supakit@it.kmitl.ac.th, ²s53070064@kmitl.ac.th, ³s53070074@kmitl.ac.th

บทคัดย่อ

การลงทุนในตลาดหลักทรัพย์มีความที่ไม่แน่นอน การตัดสินใจลงทุนจึงมีความเสี่ยง ข้อมูลที่มากพอและการวิเคราะห์ที่ดี ย่อมช่วยลดความเสี่ยงการตัดสินใจที่ผิดพลาดของนักลงทุนได้ งานวิจัยนี้นำฮิดเดนมาร์คอฟโมเดลร่วมกับราคาหลักทรัพย์ในรูปแบบกราฟแท่งเทียน ใช้ในการพยากรณ์การเปลี่ยนแปลงของหลักทรัพย์ งานวิจัยนี้ได้ออกแบบปรับปรุงวิธีการแปลงราคาของหลักทรัพย์ให้เป็นลักษณะแท่งเทียน ทั้งหมด 19 แบบ เพื่อเป็นข้อมูลให้ฮิดเดนมาร์คอฟโมเดลใช้ในการรู้จำและพยากรณ์ ซึ่งทดลองกับหลักทรัพย์ 3 กลุ่มตัวอย่างคือ SetHD, Big Cap และ Most Swing

คำสำคัญ: ฮิดเดนมาร์คอฟโมเดล, ระบบพยากรณ์กราฟแท่งเทียน

Abstract

Investment in the stock market has the uncertainty gain. Correct decision to investment and having the data or tool enough to analysis can reduce the risk. This paper presents hidden markov model forecasting the price of stock by using candlestick to be input data. This research designed 19 candlestick patterns running in the system. In experiment, the research uses three stocks from three groups in the stock exchange in Thailand, consisting setHD, big-cap and most-swing stock.

Keywords: Hidden Markov Model, Forecasting System
Candle Stick Chart

1. บทนำ

ปัจจุบันการลงทุนในตลาดหลักทรัพย์เป็นที่นิยมอย่างแพร่หลาย แม้จะมีความเสี่ยงสูง หากเกิดความสำเร็จในการลงทุน ผลตอบแทนสูงย่อมสูงตามด้วย ดังนั้นการตัดสินใจลงทุนจึงควรมีข้อมูลมากพอในการตัดสินใจ ด้วยเทคโนโลยีที่ทันสมัย ทำให้มีการบริการข้อมูลเกี่ยวกับหุ้นที่หลากหลาย เช่น ราคาหุ้นย้อนหลัง กราฟแท่งเทียน บทความวิเคราะห์จากนักวิเคราะห์หลักทรัพย์มากมาย อย่างไรก็ตามความพยายามในการพยากรณ์ราคาหลักทรัพย์เป็นสิ่งที่ยากที่จะสร้างระบบนั้นขึ้นมา การพยากรณ์ราคาของหลักทรัพย์มีพื้นฐานมาจากการวิเคราะห์ด้วยกราฟแท่งเทียน(Candlestick chart) ซึ่งถูกค้นพบขึ้นโดยนักเกร็งกำไรชาวญี่ปุ่นชื่อ Sokyu Homma ได้ทำการคิดค้นขึ้นจากพฤติกรรมราคาข้าว ปัจจุบันกราฟแท่งเทียนนิยมใช้มากในการดูลักษณะการเปลี่ยนแปลงราคา

ฮิดเดนมาร์คอฟโมเดล (Hidden Markov Model) เป็นโมเดลทางสถิติ สามารถนำมาประยุกต์ใช้กับระบบรู้จำในรูปแบบต่างๆที่หลากหลาย เช่น ระบบรู้จำเสียง ระบบรู้จำภาพ รวมถึงการพยากรณ์ เช่น การพยากรณ์สภาพอากาศ เป็นต้น ตัวอย่างงานวิจัยของ เทอดพงษ์ แดงสี และ สุพรรณ พงษ์เดช [1] ออกแบบให้ฮิดเดนมาร์คอฟโมเดลรู้จำเสียงพูดตัวเลขภาษาไทย

ผู้วิจัยจำนวนไม่น้อยทดลองสร้างระบบพยากรณ์ราคาของหลักทรัพย์ ดังเช่น งานวิจัย [2] และ [3] ใช้ Neural network ในการสร้างทำนายหลักทรัพย์ของตลาดคลิงค์โพร และใช้ Particle swarm optimization (PSO) ในการช่วยปรับค่าให้แก่ Neural network งานวิจัยดังกล่าวไม่ได้บอกชัดว่าใช้ราคาหลักทรัพย์ในลักษณะใดนำมาเป็นข้อมูลที่ป้อนให้แก่ระบบ การปรับโมเดลเพื่อรู้จำนั้นใช้ระยะเวลาในการสอน จึงประยุกต์ใช้ PSO

นำผสมผสานให้ค้นหาปรับค่าได้ผลลัพธ์เร็วขึ้น อย่างไรก็ตามในงานวิจัย [4] ได้ประยุกต์ใช้ฮิดเดนมาร์คอฟโมเดลในการทำนายหลักทรัพย์ของกลุ่มสายการบิน โดยสร้างโมเดลให้มีลักษณะ ซ้ายไปขวา (Left-right topology) กำหนดให้มีขนาดโมเดลที่ค่อนข้างจำกัดให้ hidden state มีขนาด 4 อินพุตประกอบด้วย ราคาสูงสุด ราคาต่ำสุด และราคาปิด ทางด้านเอาต์พุตใช้ ราคาปิดในวันถัดไป ผู้วิจัยเปรียบเทียบผลระหว่างกับการใช้ Neural Network อีกรงานวิจัยหนึ่ง [5] ผู้วิจัยเปรียบเทียบการใช้ฮิดเดนมาร์คอฟโมเดลนำมาทำนายราคาหลักทรัพย์ โดยผลการทดลองเปรียบเทียบความสามารถ Neural network ข้อมูลที่ป้อนให้แก่โมเดลได้แก่ ค่าการเปลี่ยนแปลงของราคาหลักทรัพย์ตั้งแต่ช่วงเปิดและปิด ราคาสูงสุดและราคาต่ำสุด

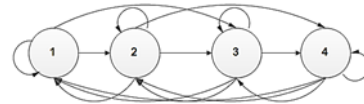
ในงานวิจัยที่กล่าวมาข้างต้นข้อมูลอินพุตที่นำมาใช้ในการพยากรณ์ราคาหลักทรัพย์เป็นข้อมูลที่อยู่ในรูปแบบของราคาปิดราคาเปิด ป้อนโดยตรงเข้าระบบ งานวิจัยนี้สนใจนำกราฟแท่งเทียนมาใช้เป็นข้อมูลอินพุตเพื่อสอนฮิดเดนมาร์คอฟโมเดลในการรู้จำและพยากรณ์ราคาหลักทรัพย์ จากการสืบค้นพบว่างานของ ธนารด ลิมสมบุญและสมาชิก [6] ได้พัฒนาระบบทำนายราคาหลักทรัพย์ด้วยฮิดเดนมาร์คอฟโมเดล เป็นระบบทำนายราคาหลักทรัพย์ โดยการนำข้อมูลมาแปลงราคาเป็นกราฟแท่งเทียนและเก็บค่าของกราฟแท่งเทียนในรูปแบบของพารามิเตอร์เพื่อใช้สำหรับการประมวลผลด้วยฮิดเดนมาร์คอฟโมเดล ผลการศึกษามีประสิทธิภาพในระดับหนึ่ง แต่ปัญหาของระบบคือส่วนของการปรับค่าพารามิเตอร์ เพื่อเพิ่มประสิทธิภาพให้กับระบบ และรูปแบบของกราฟแท่งเทียนจำนวนของแท่งเทียนและเทคนิคต่างๆ ในการวิเคราะห์อาจยังไม่เพียงพอ ซึ่งงานวิจัยนี้ได้ปรับปรุงเทคนิคและวิธีการที่ได้ใช้ในงานของ ธนารด ลิมสมบุญและสมาชิก

วัตถุประสงค์ของงานวิจัยนี้เพื่อศึกษาและสร้างการวิเคราะห์และทำนายแนวโน้มของหลักทรัพย์ โดยใช้ข้อมูลราคาหลักทรัพย์ในรูปแบบกราฟแท่งเทียน ในหัวข้อต่างๆที่จะนำเสนอมีดังนี้ ในส่วนที่ 2 นำเสนอทฤษฎีของฮิดเดนมาร์คอฟโมเดล กระบวนการฮิดเดนมาร์คอฟโมเดล ส่วนที่ 3 นำเสนอวิธีการแปลงราคาหลักทรัพย์เป็นกราฟแท่งเทียนแบบต่างๆที่ผู้วิจัยได้ออกแบบขึ้น รวมถึงขั้นตอนและวิธีการในการสอนฮิด

เดนมาร์คอฟโมเดลให้รู้จำและวิเคราะห์ทำนายแนวโน้ม ส่วนที่ 4 การกำหนดการทดลอง ส่วนที่ 5 ผลการทดลอง สรุปผลและแนวทางในการพัฒนาต่อไป

2. ฮิดเดนมาร์คอฟโมเดล

ฮิดเดนมาร์คอฟโมเดล (Hidden Markov Model) (HMM) [7][8] เป็นโมเดลทางสถิติ พัฒนามาเพื่อแบ่งกลุ่มข้อมูลอนุกรมทางเวลา หรือจัดกลุ่มของสัญญาณที่ไม่รู้จักให้อยู่ในกลุ่มเดียวกัน ฮิดเดนมาร์คอฟโมเดลได้นำมาประยุกต์ใช้ในงานเรียนรู้และจดจำข้อมูลต่างๆ โดยใช้ลำดับเหตุการณ์ที่เราสนใจ เช่น ลำดับสัญญาณเสียงพูด ลำดับของรูปภาพที่ผ่านการเข้ารหัส ลำดับของข้อมูลภูมิอากาศ เป็นข้อมูลนำเข้าให้แก่ฮิดเดนมาร์คอฟโมเดล



ภาพที่ 1: Hidden Markov Model แบบ Fully Connect topology

ส่วนประกอบของฮิดเดนมาร์คอฟโมเดลมีพารามิเตอร์ดังนี้ N คือ จำนวนของสถานะในโมเดล, T คือ ความยาวของลำดับของค่าที่ปรากฏ A คือ ค่าการกระจายความน่าจะเป็นของการเปลี่ยนสถานะกำหนดให้ $A = \{a_{ij}\}$, B คือ ค่าการกระจายความน่าจะเป็นของข้อมูลในสถานะต่างๆ ลักษณะเหมือนกับสเตตแมชชีน (state machine) เขียนเป็น $B = \{b_{ik}\}$ และ π คือ ค่าเริ่มต้นของสถานะ ดังนั้นในฮิดเดนมาร์คอฟโมเดลจึงเขียนแทนด้วย $\lambda = \{A, B, \pi\}$ มักจะเขียนแทนด้วย $\lambda = M$ ลำดับข้อมูลที่ป้อนให้แก่โมเดลจะอยู่ในรูป $O_T = \{o_1, o_2, \dots, o_T\}$

จากภาพที่ 1 แสดงโมเดลโดยเป็นลักษณะมีการเปลี่ยนสถานะได้ทั้งไปและกลับ เรียกว่า Fully connected topology ซึ่งเป็นรูปแบบที่งานวิจัยนี้ใช้ในการทดลอง

หลักการทํางานของฮิดเดนมาร์คอฟโมเดล โดยอาศัยปัญหาพื้นฐาน 3 ข้อ ต้องใช้อัลกอริทึมและวิธีการต่างๆในการแก้ปัญหา

ปัญหาที่ 1 การคำนวณหาความน่าจะเป็นของสิ่งที่สนใจเมื่อแบบจำลอง λ ได้รับลำดับของข้อมูลหนึ่งแล้ว จะมีค่าความน่าจะเป็นเท่าใด การแก้ปัญหาที่ 1 ทำได้โดยการใช้การคำนวณ

แบบไปข้างหน้า (Forward procedure) ในการคำนวณ ประกอบด้วย 3 ส่วนดังนี้

1. Initialization (การเริ่มต้น)

$$\alpha_1(i) = \pi_i b_i(O_1); \quad 1 \leq i \leq N \quad (1)$$

2. Induction (การเหนี่ยวนำ)

$$\alpha_t(i) = [\sum_{j=1}^N \alpha_{t-1}(j) a_{ij}] b_j(O_{t+1}); \quad 1 \leq t \leq T-1, \quad 1 \leq j \leq N \quad (2)$$

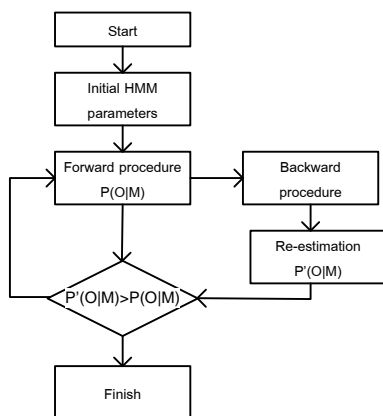
3. Termination (การสิ้นสุด)

$$P(O|\lambda) = \sum_{i=1}^N \alpha_T(i); \quad 1 \leq i \leq N \quad (3)$$

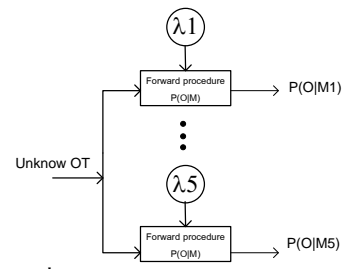
ปัญหาที่ 2 การเลือกเส้นทางของลำดับสถานะที่ดีที่สุด โมเดลนั้นไม่สามารถคาดเดาจากการสังเกตได้ จึงจำเป็นต้องหาวิธีที่ต้องคำนวณหาค่าลำดับสถานะที่ได้ค่าความน่าจะเป็นมีค่ามากที่สุด การแก้ปัญหานี้ โดยทั่วไปใช้สามารถคำนวณได้จากหลักการของ Viterbi Algorithm [7] หรือกระบวนการ Backward procedure

ปัญหาที่ 3 การปรับค่าพารามิเตอร์ของโมเดล การปรับค่าพารามิเตอร์ของโมเดลนั้น เพื่อให้ได้ค่าความน่าจะเป็นที่มีความเป็นสูงที่สุดที่เป็นไปได้ กระบวนการนี้ใช้การคำนวณเปรียบเทียบค่าความน่าจะเป็นของสถานะก่อนหน้า และสถานะถัดไป โดยใช้สมการ Baum-Welch [7] หรือ Reestimation

กระบวนการคำนวณค่าความน่าจะเป็นของฮิดเดนมาร์คอฟโมเดล [9] ซึ่งเป็นส่วนการสอนฮิดเดนมาร์คอฟโมเดลให้รู้จำ หรือแบ่งแยกสัญญาณหรือข้อมูลที่ไม่ทราบ ทำได้โดยการนำข้อมูลเหล่านั้น ซึ่งอยู่ในรูปแบบของลำดับข้อมูลที่ปรากฏ ทำการสอนให้ฮิดเดนมาร์คอฟโมเดลรู้จำ โดยมีขั้นตอนดังภาพที่ 2



ภาพที่ 2: กระบวนการสอนฮิดเดนมาร์คอฟโมเดล



ภาพที่ 3: การใช้งานฮิดเดนมาร์คอฟโมเดล

เริ่มต้น Initial HMM parameter คือการสุ่มค่าพารามิเตอร์ให้กับโมเดล พร้อมกับรับอินพุต O_T แล้วทำการหาค่าความน่าจะเป็นด้วย Forward Procedure ซึ่งจะได้ค่าความน่าจะเป็นของโมเดล $P(O|M)$ จากนั้นจึงแยกเป็นสองทาง ทางหนึ่งสร้างเป็น M' เพื่อไปคำนวณ Backward procedure และ Reestimation ปรับเพื่อให้ได้ค่าความน่าจะเป็นสูงขึ้น การคำนวณดังกล่าวเป็นการทำซ้ำ (Iterative process) และเป็น Local maxima จากนั้นจึงนำผลลัพธ์มาเปรียบเทียบกับ ค่าความน่าจะเป็นเดิม หากมีค่าความน่าจะเป็นสูงขึ้นและยังไม่ถึงเวลาที่กำหนดให้สิ้นสุดการคำนวณ ระบบก็จะวนกลับเข้าไปทำ Forward procedure, Backward procedure และ Reestimation ซ้ำจนกระทั่งสิ้นสุดตามเวลาที่กำหนด จึงเก็บค่าพารามิเตอร์ของโมเดลเพื่อใช้ในการคำนวณการรู้จำต่อไป

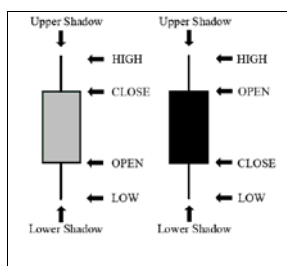
การใช้งานฮิดเดนมาร์คอฟโมเดลในการรู้จำนั้น เมื่อสอนโมเดลให้รู้จำ O_T ด้วยกระบวนการดังภาพที่ 2 เสร็จแล้วนั้น การใช้ฮิดเดนมาร์คอฟโมเดลรู้จำสร้างได้ตามโมเดลที่กำหนดให้จำในลักษณะดัง ภาพที่ 3 สมมุติมีทั้งหมด 5 โมเดลที่ผ่านการสอนมาแล้ว พร้อมกับมี O_T ที่ไม่ทราบถึงรูปแบบว่าเป็นลักษณะใด และนำ O_T นั้นป้อนเป็นอินพุตส่งให้แก่โมเดลทั้งหมด โมเดลทั้งหมดจะส่งค่าความน่าจะเป็นกลับมา โดยโมเดลใดที่ส่งค่าความน่าจะเป็นสูงสุดแสดงให้เห็นว่า O_T นั้นมีลักษณะเหมือนกับ โมเดลนั้น

3. กราฟแท่งเทียน

กราฟแท่งเทียน คือ รูปแบบของแผนภูมิแท่งเทียน ประกอบด้วย ราคาเปิด ราคาปิด ราคาต่ำสุด และราคาสูงสุด รวมไว้ในกราฟหนึ่งแท่ง เพื่อให้ผู้อ่านสามารถเข้าใจการเปลี่ยนแปลงราคาได้อย่างความรวดเร็ว ง่ายต่อการอ่านข้อมูล

รูปแบบของกราฟแท่งเทียนจึงเป็นที่นิยมอย่างยิ่งในปัจจุบัน และเหมาะสมแก่นักลงทุนโดยทั่วไป เพื่อใช้ในการวิเคราะห์ ข้อมูลหุ้นและช่วยในการตัดสินใจให้นักลงทุน โครงสร้างของกราฟแท่งเทียน [10] แสดงดังภาพที่ 4 ประกอบด้วย 4 ส่วนหลัก คือ ราคาเปิด ราคาปิด ราคาต่ำสุด และ ราคาสูงสุด

ส่วนประกอบต่างๆที่เกี่ยวข้อง 1. แท่งตรงกลางของกราฟแท่งเทียน เรียกว่า แท่งเทียน (Real Body), 2. ไม้เทียนทางบนของกราฟแท่งเทียน เรียกว่า (Upper Shadow) และ 3. ไม้เทียนทางล่างของกราฟแท่งเทียน เรียกว่า (Lower Shadow) การเปลี่ยนแปลงกราฟแท่งเทียนในกรณีพื้นฐานที่เกิดขึ้น ได้แก่ ในกรณีที่ราคาปิด (Close) สูงกว่าราคาเปิด (Open) แท่งเทียนจะเป็นสีขาวหรือสีเขียว เรียกว่า White Candlestick เป็นการแสดงถึงแนวโน้มที่ดีคือ ราคามีการปรับตัวขึ้น อีกรูปแบบหนึ่งในกรณีที่ ราคาปิดต่ำกว่าราคาเปิดแท่งเทียนจะเป็นสีดำหรือสีแดง เรียกว่า Black Candlestick แสดงถึงแนวโน้มที่ไม่ดีคือ ราคามีการปรับตัวลง และในกรณีที่ราคาปิดเท่ากับราคาเปิด โดยราคาดังนั้นอาจเป็นราคาต่ำสุดหรือสูงสุดก็ตามแต่แท่งเทียนจะมีลักษณะเป็นเส้นขีดขวาง เรียกว่า Doji Stick เป็นการแสดงลักษณะที่เป็นกลาง ในกรณีที่เกิดขึ้นในช่วงขาขึ้น อาจแสดงถึงจุดกลับตัวเป็นขาลง หรือในกรณีที่ในช่วงขาลง อาจแสดงถึงจุดกลับตัวเป็นช่วงขาขึ้นได้



ภาพที่ 4: โครงสร้างและส่วนประกอบของกราฟแท่งเทียน [11] [12]

4. การออกแบบให้ HMM รู้จำกราฟแท่งเทียน

งานวิจัยนี้ทำการรวบรวมข้อมูลราคาหุ้นย้อนหลัง จากตลาดหลักทรัพย์แห่งประเทศไทย เพื่อนำมาทำการวิเคราะห์และพัฒนาระบบ โดยมีขั้นตอนการดำเนินงาน เริ่มต้นจาก 1) การนำราคาย้อนหลังจากตลาดหลักทรัพย์ โดยการดึงข้อมูลผ่านทางเว็บไซต์ www.settrade.com ซึ่งจะมีราคาเปิดราคาปิดย้อนหลังเป็นระยะเวลา 6 เดือน 2) นำราคาเปิดราคาปิดในแต่ละวันของ

หลักทรัพย์มาแปลงเป็นกราฟแท่งเทียน โดยกำหนดรูปแบบแท่งเทียนทั้งหมด 19 แบบ จากนั้น 3) ข้อมูลราคาที่แปลงเป็นแท่งเทียนแล้วจะนำเป็นข้อมูลป้อนให้แก่ฮิดเดนมาร์คอฟโมเดลสร้างเป็นโมเดลตามลักษณะแท่งเทียนทั้งหมด 19 โมเดล และสุดท้าย 4) นำราคาการเปลี่ยนแปลงของหลักทรัพย์ความยาว 8 วัน และ 10 วันสุดท้ายมาเป็นอินพุตป้อนให้แก่โมเดลทั้ง 19 เพื่อพยากรณ์ว่าในวันถัดไปจะได้กราฟแท่งเทียนในลักษณะใด

ตารางที่ 1: รูปแบบ(1-10) ของแท่งเทียนที่ใช้ในการวิเคราะห์ข้อมูล

ลำดับ	ลักษณะแท่งเทียน	ลำดับ	ลักษณะแท่งเทียน
1		6	
Bullish Belt Hold Line green		Bearish Belt Hold Line red	
2		7	
Long Body Long Core green		Body Long Core red	
3		8	
Short Body Short Core green		Short Body Short Core red	
4		9	
Short Body Long Core green		Short Body Long Core red	
5		10	
Long Body Short Core green		Long Body Short Core red	

ตารางที่ 2: รูปแบบ(11-18) ของแท่งเทียนที่ใช้ในการวิเคราะห์ข้อมูล

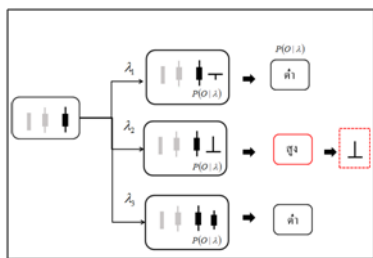
ลำดับ	ลักษณะแท่งเทียน	ลำดับ	ลักษณะแท่งเทียน
11		16	
Four Price Doji		Long-Legged Doji both short	
12		17	
Long - Dragonfly Doji		Long-Legged Doji both long	
13		18	
Short - Dragonfly Doji		Long-Legged Doji high long	
14		19	
Short Gravestone Doji		Long-Legged Doji high short	
15			
Long Gravestone Doji			

การรวบรวมข้อมูลหุ้น เป็นส่วนของการรับข้อมูล โดยระบบจะทำการจัดเก็บข้อมูลหุ้นย้อนหลังแล้วทำการแปลงข้อมูลหุ้นที่ได้ เป็นกราฟแท่งเทียน รวมถึงการแปลงข้อมูลให้อยู่ในรูปแบบของชุดข้อมูลสำหรับฮิดเดนมาร์คอฟโมเดล หรือ O_T

ลักษณะของแท่งเทียนในงานวิจัยนี้กำหนดทั้งหมด 19 แบบ แสดงดังตารางที่ 1 ถึง 2 แท่งเทียนที่แสดงหลักทรัพย์ในสถานะราคาเพิ่มขึ้นได้แก่หมายเลข 1 2 3 4 5 และปรับตัวลดลงได้แก่หมายเลข 6 7 8 9 10

หลักการวิเคราะห์และทำนายแนวโน้มของหลักทรัพย์นั้น เมื่อนำข้อมูลราคาหลักทรัพย์มาแปลงเป็นแท่งเทียนที่เข้ารหัสสัญลักษณ์เป็นตัวเลข 1 ถึง 19 (ทั้งหมด 19 แบบ) จะทำการสร้างฮิดเดนมาร์คอฟโมเดลทั้งหมด 19 โมเดล ลักษณะตัวอย่างดังภาพที่ 5 (ตัวอย่างสร้างสามโมเดล) โมเดลทั้งหมดจะได้รับการสอนตามลักษณะแท่งเทียน โดยจัดแบ่งความยาวของ O_T ที่ 8 และ 10 แท่งเทียน เปรียบเสมือนกับราคาของหลักทรัพย์ที่เปลี่ยนแปลงในช่วงเวลา 8 วัน และ 10 วัน

ดังนั้นการพยากรณ์จะใช้ข้อมูลลักษณะแท่งเทียนจำนวน 8 วันและ 10 วันก่อนหน้า มาทำการวิเคราะห์และทำนายลักษณะแท่งเทียนในวันถัดไป



ภาพที่ 5: หลักการพยากรณ์ลักษณะของแท่งเทียน

5. การทดลองและผลการทดลอง

การทดลองในช่วงวันที่ 13 ถึง 17 มกราคม 2557 ซึ่งเป็นช่วงที่ราคาหลักทรัพย์มีความไม่แน่นอนค่อนข้างสูง และต่างชาติมีการขายหลักทรัพย์ออกในปริมาณมาก พารามิเตอร์ที่เกี่ยวข้องประกอบด้วย ลักษณะของฮิดเดนมาร์คอฟโมเดลเป็นแบบ Fully connected มีขนาดของ A คือ 30x30 และ 35x35 B คือ 30x19 และ 35x19 ความยาวของ O_T กำหนดที่ 8 และ 10 (8วัน

และ 10 วัน) การปรับพารามิเตอร์ทำการเปรียบเทียบผลลัพธ์ โดยเลือกใช้การปรับพารามิเตอร์ลักษณะ Heuristic Random search การสุ่มสร้างพารามิเตอร์ และ Baum-Welch Algorithm โดยกำหนดจำนวนการสุ่มที่ 100 ครั้ง และ 1,000 ครั้ง ตามลำดับ โปรแกรมทั้งหมดเขียนด้วยภาษา Java (Netbean) ข้อมูลหลักทรัพย์ที่นำมาพิจารณาใช้ข้อมูลราคาเปลี่ยนแปลงย้อนหลัง 6 เดือนของหลักทรัพย์ 3 กลุ่มประกอบด้วย TISCO, LPN, BICG, KTB, TRUE, และ QH

ตารางที่ 3: ผลการทดลองการพยากรณ์แนวโน้มแท่งเทียน

Observation length, Size of Matrix A	Heuristic Search		Baum-Welch Method	
	ทำนาย	เปอร์เซ็นต์	ทำนาย	เปอร์เซ็นต์
Obs 30x30, 8,	13/35	37%	25/35	72%
Obs 30x30, 10,	15/35	43%	20/35	57%
Obs 35x35, 8,	17/35	49%	14/35	40%
Obs 35x35, 10,	8/35	23%	25/35	72%

จากการทดลอง ผลการทดลองที่เกิดขึ้น แสดงตารางเปรียบเทียบการวัดความถูกต้องในกรณีวัดแท่งเทียนในกลุ่มเดียวกัน (แนวโน้มไปในทางเดียวกัน) เช่น ผลการวิเคราะห์และทำนายแสดงลักษณะราคาเพิ่มขึ้น แสดงระหว่างหมายเลข 1-5

วิธีการสุ่ม Heuristic search มีค่าความถูกต้อง 37% สำหรับ $A = 30 \times 30$ ชุดข้อมูล 8, มีค่าความถูกต้อง 43% สำหรับชุดข้อมูลขนาด 10

วิธีการสุ่ม Heuristic search มีค่าความถูกต้อง 49% สำหรับ $A = 35 \times 35$ ชุดข้อมูล 8, มีค่าความถูกต้อง 23% สำหรับชุดข้อมูลขนาด 10

วิธีการปรับพารามิเตอร์ Baum-Welch มีค่าความถูกต้อง 72% สำหรับ $A = 30 \times 30$ ชุดข้อมูล 8, มีค่าความถูกต้อง 57% สำหรับชุดข้อมูลขนาด 10

วิธีการปรับพารามิเตอร์ Baum-Welch มีค่าความถูกต้อง 40% สำหรับ $A = 35 \times 35$ ชุดข้อมูล 8, มีค่าความถูกต้อง 72% สำหรับชุดข้อมูลขนาด 10

จากผลการทดลองข้างต้น แสดงให้เห็นถึงประสิทธิภาพการ สอนโมเดลแต่ละวิธี โดยวิธี Baum-Welch มีประสิทธิภาพการ สอนที่ดีกว่าแบบวิธีสุ่ม Heuristic Search เนื่องจากวิธีการสุ่ม Heuristic Search คือการปรับค่าพารามิเตอร์ ด้วยวิธีการสุ่มค่า อย่างเดียวเท่านั้น แต่สำหรับวิธีการ Baum-Welch เป็นการปรับ ค่าพารามิเตอร์ อ้างอิงจากการปรับค่าครั้งก่อนหน้า เพื่อให้ค่า ของโมเดลมีประสิทธิภาพที่ดีขึ้น การปรับแต่ละครั้งค่าที่ได้จะมี ประสิทธิภาพดีขึ้น และปรับจนกว่าค่าจะไม่เกิดการ เปลี่ยนแปลง

6. บทสรุปและคำวิจารณ์

ระบบวิเคราะห์และพยากรณ์แนวโน้มหลักทรัพย์ โดยใช้ ทฤษฎีฮิดเดนมาร์คอฟโมเดล ในการศึกษาพิจารณากราฟ แท่งเทียนเป็นอินพุต เพียงอย่างเดียว ซึ่งผลการทดลองมีความ ถูกต้องในระดับหนึ่ง สะท้อนให้เห็นถึงความจำเป็นต้องใช้ ปัจจัยอื่นๆเข้ามาใส่เป็นอินพุตในให้ฮิดเดนมาร์คอฟโมเดลได้ เรียนรู้เพิ่ม เพราะปัจจัยการเปลี่ยนแปลงของหลักทรัพย์ ไม่ เหมือนการเปลี่ยนแปลงของภูมิอากาศ เพราะหลักทรัพย์มี อิทธิพลที่เกี่ยวข้องหลากหลาย เช่นเงินทุนต่างชาติ อัตรา แลกเปลี่ยนทางการเงิน และปัจจัยทางการเมือง เป็นต้น อย่างไรก็ตามหากได้ออกแบบฮิดเดนมาร์คอฟให้ใส่ปัจจัยที่เกี่ยวข้องได้ หลากหลายและมากยิ่งขึ้น ผลความคลาดเคลื่อนของการ พยากรณ์ย่อมน้อยลงตามไปด้วยเช่นกัน

เอกสารอ้างอิง

- [1] เทอดพงษ์ แดงสี และ สุวัฒน์ พงษ์เดช. “การรู้จำ เสียงพูดตัวเลข ภาษาไทย 0-9 แบบขึ้นกับผู้พูด โดยวิธีฮิดเดนมาร์คอฟโมเดล” *ปริญญานิพนธ์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*, 2539
- [2] Boo Junyou, “Stock Price forecasting using PSO-trained neural networks”, *IEEE congress on Evolutionary Computation*, pp.2879,2885, 25-28 Sept. 2007
- [3] Lu Bo, “A stock prediction method based on PSO and BP hybrid algorithm”, *E -Business and E -Government (ICEE), 2011 International Conference on* , pp.1,4, 6-8 May 2011

- [4] Hassan, M.R.; Nath, B., “Stock market forecasting using hidden Markov model: a new approach”, *Intelligent Systems Design and Applications, 2005. ISDA '05. Proceedings. 5th International Conference on*, pp.192,196, 8-10 Sept. 2005
- [5] Gupta, A.; Dhingra, B., “Stock market prediction using Hidden Markov Models,” *Students Conference on Engineering and Systems (SCES)*, pp.1,4, 16-18 March 2012
- [6] ชนารถ ลิมสมบุญชัย, ธนะศักดิ์ ลิขิตยั้งยืน และ ธนา โชคพรชูดิรัก. “ระบบทำนายราคาหุ้นด้วยฮิดเดนมาร์คอฟโมเดล”, *ปริญญานิพนธ์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง*, 2550.
- [7] Lawrence R. Rabiner. “A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition,” *Proceeding of the IEEE*, Vol. 77, no. 8, Feb 1989.
- [8] Anders Meng. “An introduction to Markov and Hidden Markov Models”, *Online book download at www2.imm.dtu.dk/pubdb/views/edoc_download.php/3313/pdf/imm3313.pdf*, Oct 2003.
- [9] Philip Jackson. “HMM tutorial 3.”, http://personal.ee.surrey.ac.uk/Personal/P.Jackson/tutorial/hmm_tut3.
- [10] ประจวบ วงษ์นิยม, “พินิจหุ้นด้วยกราฟแท่งเทียน”, *สำนักพิมพ์กรุงเทพธุรกิจ Bizbook*, 2553
- [11] Steve Nison, “Japanese Candlestick Charting Techniques 2nd,” *Prentice Hall Press*, p.299, 2001

การวิเคราะห์ข้อมูลและแสดงผลข้อมูลเชิงพื้นที่และเวลาบนกูเกิลเอิร์ธ

Spatial-Temporal Analysis and Visualization of Twitter Data on Google Earth

อิสราภรณ์ วิทธีวราพันธ์ (Isaraporn Vithyaviranont)¹ และวีระ เหมือนสิน (Veera Muangsin)²

ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

¹Isaraporn.V@Student.chula.ac.th, ²Veera.M@chula.ac.th

บทคัดย่อ

ทวิตเตอร์เป็นสื่อสังคมแบบไมโครบล็อกและส่งข้อความที่ได้รับความนิยมที่สุดแห่งหนึ่ง ด้วย Streaming API และบริการระบุพิกัดบนข้อความทำให้ทวิตเตอร์เป็นแหล่งข้อมูลขนาดใหญ่สำหรับการวิเคราะห์สื่อสังคมในเชิงพื้นที่และเวลา บทความนี้นำเสนอการออกแบบและพัฒนาระบบแอปพลิเคชันบนเว็บซึ่งเรียกว่า World of Tweeties (WOT) เพื่อใช้เป็นเครื่องมือในการวิเคราะห์ข้อมูลเชิงพื้นที่และเวลาของข้อความทวิตเตอร์ ผู้ใช้งานสามารถค้นหาข้อความทวิตด้วยคีย์เวิร์ด, ชื่อผู้ใช้ และช่วงเวลา รวมถึงการค้นหาด้วยหลายคีย์เวิร์ดเพื่อการเปรียบเทียบข้อมูล สำหรับการวิเคราะห์จะมีทั้งเชิงพื้นที่ซึ่งแสดงข้อมูลบนกูเกิลเอิร์ธ และเชิงเวลาจะถูกนำเสนอในรูปแบบของกราฟจำนวนทวีตบนแกนเวลา นอกจากนี้ระบบยังแสดงผลจากการคำนวณทางสถิติและข้อมูลเชิงพื้นที่แบบตัวเลขอีกด้วย ผลลัพธ์ที่ได้สามารถจัดเก็บในคลังข้อมูลเพื่อให้เรียกดูได้ในภายหลัง การทดลองวิเคราะห์ทวิตในประเทศไทยให้ผลลัพธ์ที่น่าสนใจและมีประโยชน์ในหลายแนวทาง เช่น ในการสำรวจด้านการเมือง ติดตามสภาพการจราจร สภาพอากาศ หรือวิเคราะห์การตลาด เป็นต้น

คำสำคัญ: ทวิตเตอร์, กูเกิลเอิร์ธ, การวิเคราะห์เชิงพื้นที่, การวิเคราะห์สื่อสังคม

Abstract

Twitter is a popular social media system with micro-blogging and messaging features. With its Streaming API

and geolocation attached data, Twitter provides a huge data source for spatial-temporal analysis on social media. This paper presents the design and development of a Web-based application called "World of Tweeties (WOT)". It is an on-line spatial-temporal analysis tool for Twitter data. The user can search the tweets by keyword, username, and time. Searching with multiple keywords for comparison is supported. The Google Earth plug-in provides the geographic user interface for identifying the area of interest and data visualization. Temporal analysis is supported with a time-frequency graph. The application also performs statistic and spatial analysis. The results can be archived for later examination. Experiments on the tweets originated in Thailand show interesting results and the potentials of the system in many areas such as political survey, traffic monitoring, weather monitoring, brand analysis, etc.

Keyword: Twitter, Google Earth, spatial analysis, social media analysis.

1. บทนำ

ทุกวันนี้ ผู้คนนับล้านกำลังร่วมกันสร้างข้อมูลดิจิทัลผ่านสื่อสังคมออนไลน์ต่างๆ เช่น Facebook, Twitter, Instagram, YouTube, Google+ ฯลฯ โดยมีการติดต่อสื่อสาร การเผยแพร่ข้อมูลผ่านสถานะในเพจ การแบ่งปันรูปภาพและวิดีโอ การให้

คะแนนหรือรีวิวลินค้าและบริการ ฯลฯ ซึ่งเรียกกันว่าเป็นวัฒนธรรมแบบเว็บ 2.0

ทวีตเตอร์เป็นหนึ่งในบรรดาสื่อสังคมออนไลน์ที่ถูกศึกษาวิจัยมากที่สุด ทวีตเตอร์เป็นระบบไมโครบล็อกและบริการส่งข้อความซึ่งก่อตั้งในปี 2006 [1] ผู้ใช้งานสามารถเผยแพร่ข้อมูลผ่านข้อความ ขนาดไม่เกิน 140 ตัวอักษร ซึ่งเรียกว่า “ทวีต (Tweet)” บนหน้าเพจของตัวเอง ซึ่งอนุญาตให้ผู้ใช้คนอื่นๆ อ่านได้ และผู้ใช้แต่ละคนสามารถ “ติดตาม (Follow)” ผู้ใช้คนอื่นๆ เพื่อให้ทวีตจากผู้ใช้คนนั้นมาปรากฏบนหน้าเพจของตน ซึ่งกลไกนี้นำไปสู่การเผยแพร่ข่าวสารและการสนทนาในกลุ่มผู้ใช้งานที่ซ้อนทับกันได้ นอกจากนี้ ผู้ใช้ยังสามารถ “ส่งต่อ (Retweet)” ทวีตของคนอื่น ซึ่งจะแสดงโพสต์นั้นๆ บนหน้าเพจของตัวเองได้อีกด้วย ซึ่งทำให้เกิดการกระจายข้อมูลจากกลุ่มหนึ่งไปยังอีกกลุ่มหนึ่งต่อไปเรื่อยๆ ทุกวันนี้ทวีตเตอร์มีจำนวนผู้ใช้งานแต่ละวันมากกว่า 100 ล้านคน ซึ่งส่งข้อมูลไม่ต่ำกว่า 500 ล้านทวีตต่อวัน [2] สำหรับในประเทศไทยมีผู้ใช้งานจริงมีประมาณ 200,000 คนต่อวัน ซึ่งส่งกว่า 1.1 ล้านทวีตต่อวัน

มีนักวิจัยจำนวนมากได้ศึกษาค้นคว้าและวิเคราะห์ข้อมูลของทวีตเตอร์ เพื่อใช้ประโยชน์ในหลากหลายแง่มุม เช่น การตรวจสอบความพึงพอใจของผู้ใช้สินค้าและบริการต่างๆ การวัดกระแสสังคมว่ากำลังติดตามเรื่องใดมากขึ้นหรือน้อยลง การติดตามทวีตในหัวข้อเฉพาะเรื่อง การประเมินอารมณ์ของผู้ใช้จากข้อความที่โพสต์ว่าเป็นแง่บวกหรือแง่ลบ เหล่านี้เป็นการวิเคราะห์ส่วนของเนื้อหาความในเชิงเวลา

ข้อมูลที่มนุษย์สร้างขึ้นส่วนใหญ่จะเกี่ยวข้องกับกิจกรรมและความสนใจของมนุษย์ ซึ่งมักมีความเกี่ยวข้องกับเวลาและสถานที่ เวลานี้คอมพิวเตอร์พกพามีความสามารถในการระบุพิกัดตำแหน่งจาก GPS จึงเป็นไปได้ที่จะแนบข้อมูลตำแหน่งที่ตนอยู่ไปกับข้อมูลที่ส่งด้วย ซึ่งทำให้เกิดข้อมูลในเชิงพื้นที่และเวลามากขึ้น ดังนั้น หากเราสามารถมองเห็นกิจกรรมออนไลน์ที่เกิดขึ้นในแต่ละพื้นที่อย่างต่อเนื่องได้ ก็จะสามารถมองเห็นความสนใจที่เปลี่ยนไปของผู้คนในบริเวณนั้นๆ ได้เช่นกัน จึงทำ

ให้ผู้ใช้ข้สนใจในเรื่องการวิเคราะห์ข้อมูลในเชิงพื้นที่และเวลาของข้อมูลที่ได้รับจากเครือข่ายสังคม

ทวีตเตอร์มีความสามารถในการแนบพิกัดตำแหน่งไปกับข้อความได้ ทั้งยังมีชุดคำสั่ง Streaming API ที่ทำให้ผู้พัฒนาซอฟต์แวร์สามารถดึงข้อความทวีตจากทวีตเตอร์แบบ real-time จึงเป็นสื่อที่เหมาะสมสำหรับการทำวิจัย ซึ่งในเริ่มแรกมีนักวิจัยได้พัฒนาเครื่องมือเพื่อนำข้อความทวีตมาป้อนกับกฎเกิลเอิร์ธ ซึ่งทำให้ผู้ใช้สามารถดูทวีตในพื้นที่รอบโลกได้ [5]

บทความนี้นำเสนอการปรับปรุงโครงงาน [5] เพื่อเพิ่มความสามารถด้านการค้นหา การวิเคราะห์ การแสดงผล และการจัดเก็บข้อมูล เนื้อหาของบทความประกอบด้วย วรรณกรรมที่เกี่ยวข้อง วิธีการดำเนินงาน เช่น การเข้าถึงข้อความทวีตจากทวีตเตอร์ การวิเคราะห์เชิงพื้นที่และเวลา และการแสดงผลด้วยแผนที่และกราฟ และผลการทดลองที่ได้จากการค้นหาด้วยคีย์เวิร์ดต่างๆ ในงานวิจัยได้เรียกกระบวนนี้ว่า “World of Tweets” หรือ “WOT”

2. งานวิจัยที่เกี่ยวข้อง

2.1 การทำเหมืองข้อความและการค้นคืนสารสนเทศ

การวิเคราะห์ข้อมูลบนสื่อสังคมออนไลน์ได้รับความนิยมมากขึ้นตามการขยายตัวและความนิยมของสื่อสังคมออนไลน์ที่ผู้ใช้งานต่างให้ความคิดเห็นเกี่ยวกับเรื่องรอบตัวบนสื่อเหล่านี้ ทำให้บริษัทที่ให้บริการและจำหน่ายสินค้าต่างๆ ต้องสำรวจข้อมูลเกี่ยวกับตนเองและคู่แข่งจากสื่อสังคมออนไลน์อยู่เสมอ

หนึ่งในจุดเด่นของทวีตเตอร์ก็คือชุดคำสั่งที่เปิดกว้างให้นักวิจัยสามารถเข้าถึงข้อมูลของทวีตเตอร์ได้โดยง่าย ซึ่งข้อมูลสตรีมมิ่ง (Streaming data) ที่ไหลเข้ามานั้นสามารถใช้เป็น crowd-sourced sensing [7] หรือเหมืองข้อมูลได้ ดังเช่น Son Doan และคณะ [6] ได้ศึกษาการติดตามโรคไข้หวัดใหญ่ผ่านทางทวีตเตอร์โดยใช้คีย์เวิร์ดเกี่ยวกับอาการของโรค และใช้ศัพท์ที่คนทั่วไปใช้กันเป็นตัวกรองทวีต Hotstream [4] เป็นแอปพลิเคชันสำหรับตรวจสอบข่าวที่ทันต่อเหตุการณ์โดยการรวบรวม จัดกลุ่ม เรียงลำดับ และติดตามข้อความ

ทวิตเตอร์เป็นช่องทางการสื่อสารที่นิยมอย่างมากในบางประเทศ ดังนั้น จึงมีการประยุกต์ใช้ทวิตเตอร์ในการเตือนภัยทางธรรมชาติอย่างภัยสึนามิด้วยระบบ Twitter Early Warning ซึ่งถูกพัฒนาและนำมาทดลองใช้ในประเทศอินโดนีเซีย [8] ทวิตเตอร์ยังถูกทดลองใช้ในการทำนายผลการเลือกตั้งในประเทศสเปนเมื่อปีค.ศ. 2012 โดยใช้เครื่องมือชื่อ Taratweet ซึ่งนับจำนวนทวิตที่เกี่ยวข้องกับพรรคการเมือง [3] อย่างไรก็ตาม ผลการทดลองค่อนข้างแตกต่างจากผลการเลือกตั้งจริง ซึ่งสาเหตุหนึ่งก็คือระบบไม่สามารถแยกได้ว่าข้อความที่พูดถึงพรรคการเมืองนั้นๆ ถูกพูดถึงไปในทางที่ดีหรือไม่ดี การตรวจสอบว่าข้อความนั้นๆ กำลังสื่อถึงความรู้สึกหรืออารมณ์แบบใดจึงเป็นอีกเรื่องที่น่าสนใจ จึงมีงานวิจัยที่ได้แบ่งแยกอารมณ์ของข้อความออกเป็น 6 แบบ โดยการแยกแยะแบบขั้นลำดับและได้ทำการทดลองผ่านชุดข้อมูลจากการแข่งขันฟุตบอล Brazilian Soccer League [10]

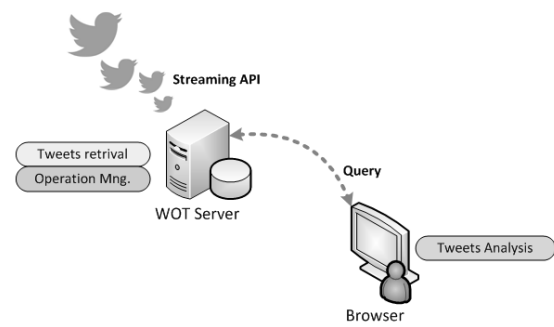
2.2 ข้อมูลสารสนเทศเชิงพื้นที่และเวลา

ทวิตเตอร์ได้เพิ่มความสามารถในการระบุพิกัดละติจูด ลองจิจูดให้กับแต่ละข้อความในเดือนสิงหาคมปี 2009 [9] ทำให้สามารถนำมาวิเคราะห์ในเชิงพื้นที่ได้ เช่น งานวิจัยของ Sakaki และคณะที่ได้นำ Kalman filter มาช่วยกรองข่าวสารจากทวิตและประมาณค่าหาจุดกำเนิดของการเกิดสึนามิในญี่ปุ่น [7] แต่ถึงแม้ว่าบริการระบุตำแหน่งจะถูกนำมาใช้บ่อยในทุกวันนี้ แต่ผู้ใช้งานส่วนใหญ่ก็ยังไม่อนุญาตให้แสดงพิกัดของตนเองในทวิต เนื่องจากประเด็นเรื่องความเป็นส่วนตัว ซึ่งปัญหานี้ทำให้มีนักวิจัยหาวิธีระบุตำแหน่งพิกัดของผู้ใช้งานโดยอาศัยข้อมูลจากข้อความในทวิตและการสนทนากับผู้ใช้งานอื่น [11]

3. วิธีการดำเนินการวิจัย

ในส่วนนี้จะกล่าวถึงการออกแบบ ขั้นตอนการทำงาน และอัลกอริทึมที่ใช้ในการวิเคราะห์สตรีมมิ่งทวิตของระบบ WOT

ระบบถูกออกแบบการทำงานให้เป็นแบบไคลเอนต์เซิร์ฟเวอร์ซึ่งประกอบไปด้วย 3 ส่วนหลักๆ ได้แก่ ส่วนค้นคืนทวิต ส่วนจัดการการทำงาน และส่วนวิเคราะห์ทวิต ดังภาพที่ 1



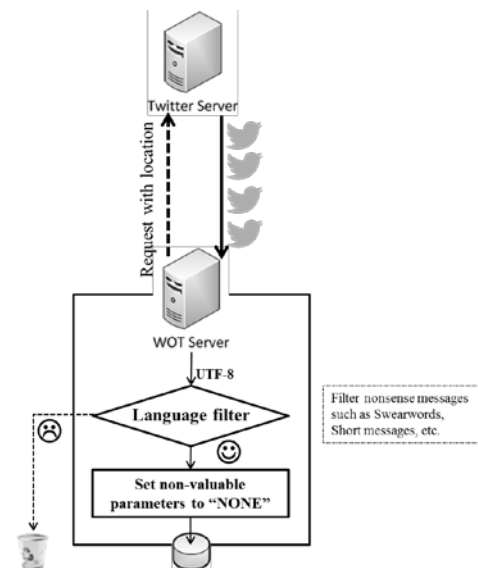
ภาพที่ 1: ภาพรวมของระบบ

3.1 ส่วนค้นคืนทวิต (Tweet Retrieval)

เพื่อร้องขอข้อมูลทวิตจากทวิตเตอร์เซิร์ฟเวอร์ โปรแกรมในฝั่งเซิร์ฟเวอร์จะทำการติดต่อทวิตเตอร์ผ่านการเรียก Public API ซึ่งเป็นชุดคำสั่งหนึ่งใน Streaming API โดยชุดคำสั่งนี้อนุญาตให้กำหนดขอรับเฉพาะทวิตที่อยู่ในพื้นที่ตามพิกัดที่กำหนดได้ [12] แต่ละทวิตจะประกอบไปด้วยข้อความ เวลาที่ถูกโพสต์ และพิกัดตำแหน่ง จากนั้นทวิตที่ได้รับจะถูกกรองเพื่อคัดทวิตที่มีค่าอันไม่พึงประสงค์ออกไป ตามภาพที่ 2

3.2 ส่วนจัดการการทำงาน (Operation Manager)

ส่วนจัดการการทำงานนี้ถือเป็นส่วนกลางของระบบซึ่งการตอบสนองจะขึ้นอยู่กับสิ่งที่ร้องขอจากผู้ใช้งาน เมื่อเซิร์ฟเวอร์ได้รับจะทำการจ่ายงานไปยังโมดูลอื่นๆ ที่เกี่ยวข้อง โดยการร้องขอสามารถแบ่งออกได้เป็น 2 คำสั่งหลัก ดังนี้



ภาพที่ 2: การทำงานของส่วนค้นคืนทวิต

1) ค้นหา: ระบบรองรับการค้นหา 3 รูปแบบ คือ

- All tweets: การรื้อขอข้อความทั้งหมดที่ถูกจัดเก็บ
- By keyword: ดึงข้อความทั้งหมดที่มีคีย์เวิร์ดนั้น
- By username: ดึงข้อความทั้งหมดของผู้ใช้งานนั้นๆ

ทั้งนี้การค้นหาจะถูกจำกัดด้วยขอบเขตการมองเห็นใน

หน้าต่างกูเกิลเอิร์ธและช่วงเวลาที่กำหนดด้วย

2) การจัดเก็บถาวร (Archive): ใช้สำรองผลการค้นหา

ทวิตเพื่อใช้เก็บข้อมูลที่น่าสนใจซึ่งต้องการเรียกดูในภายหลัง

ทั้งนี้เพื่อหลีกเลี่ยงการถูกลบข้อมูลสำหรับประหยัดเนื้อที่

- การบันทึก (Save): จัดเก็บคีย์เวิร์ดที่ใช้ค้นหาและ

ทวิตจากที่ได้จากการค้นหาแยกไว้ในฐานข้อมูลถาวร

- การเรียกข้อมูล (Load): การค้นคืนทวิตที่ถูกจัดเก็บ

ไว้ในฐานข้อมูลถาวร

- ลบบันทึก (Delete): ลบข้อมูลที่จัดเก็บเอาไว้

3.3 ส่วนวิเคราะห์ทวิต (Tweet Analysis)

โมดูลนี้จะเริ่มทำงานเมื่อรายการทวิตจากการค้นหาทางฝั่ง

เซิร์ฟเวอร์ได้ถูกส่งคืนกลับมาที่ไคลเอนต์แล้ว โดยภายในโมดูล

นี้จะมีโมดูลย่อยที่ช่วยในการวิเคราะห์อยู่ 3 โมดูลคือ กูเกิลเอิร์ธ

โมดูล, กราฟโมดูล และส่วนการวิเคราะห์

สำหรับการแสดงทวิตเป็นหมุดบนกูเกิลเอิร์ธ และแสดง

บอลลูกข้อความขึ้นเมื่อคลิกที่หมุด ดังภาพที่ 3 ทำได้โดยการ

สร้างไฟล์ KML (Keyhole Markup Language) ที่เก็บ

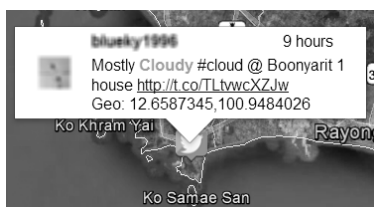
ตำแหน่งและข้อความของแต่ละทวิต

ในส่วนของกราฟ จะแสดงถึงจำนวนทวิตของคีย์เวิร์ดนั้น

บนแกน y เทียบกับเวลาในแกน x มีหน่วยเป็นวันหรือชั่วโมง

เรื่องที่ได้รับความสะดวกจะเพิ่มขึ้นอย่างกะทันหันจึงมีลักษณะ

กราฟความถี่ต่อเวลาสูงขึ้นอย่างรวดเร็ว



ภาพที่ 3: ตัวอย่างการแสดงผลทวิตบนหน้าต่างกูเกิลเอิร์ธ

การวิเคราะห์ทวิตถูกออกแบบมาเพื่อทำเหมืองข้อมูลของ

ชุดข้อความซึ่งจะได้ความรู้ในเชิงพื้นที่ออกมา เช่น ความเร็วใน

การแพร่กระจายของข่าวสารระหว่างทวิตที่เก่าที่สุดและใหม่

ที่สุด พื้นที่และความหนาแน่นของทวิตที่เกิดขึ้น ซึ่งในขั้นนี้

สมการ Havensine ถูกนำมาใช้เพื่อหาระยะทางระหว่างพิกัด

ละติจูดลองจิจูดสองจุด โดยสมมติว่าโลกเกือบจะเป็นทรงกลม

สมการนี้เป็นหนึ่งในวิธีการค้นหาระยะทางที่สำคัญในระบบนำ

ร่องซึ่งจะให้ระยะทางที่สั้นที่สุดบนพื้นผิวทรงกลม ตามสมการ

ที่ (1) (2) (3) และในส่วนของขนาดพื้นที่จะอาศัยระยะทางที่วัด

ได้จากสมการนี้ในการคำนวณ

$$d = rhavensin^{-1}(h) \quad (1)$$

$$h = haversin(\phi_2 - \phi_1) + \cos(\phi_1) \cos(\phi_2) haversin(\lambda_2 - \lambda_1) \quad (2)$$

$$havarsin\theta = \sin^2(\theta/2) \quad (3)$$

โดย ϕ แทนพิกัดละติจูด

λ แทนพิกัดลองจิจูด

d แทนระยะทางระหว่างพิกัดสองจุด

r แทนรัศมีของโลก, $\sim 6,371$ กิโลเมตร

4. ผลการดำเนินงาน

ขอบเขตการทดลองถูกจำกัดในพื้นที่ครอบคลุมประเทศ

ไทยที่ละติจูดระหว่าง 5.610000 ถึง 20.463181 และลองจิจูด

ระหว่าง 97.345619 ถึง 105.639381 ทางฝั่งของไคลเอนต์ได้

ให้ผู้ทดลองทำการค้นหาข้อมูลที่สนใจจากระบบและดูผลลัพธ์

บนบราวเซอร์ โดยแต่ละข้อความที่เกี่ยวข้องจะถูกจัดแสดงเป็น

หมุดบนกูเกิลเอิร์ธเพื่อแสดงการกระจายตัวของข้อมูลในแต่ละ

พื้นที่ และแสดงผลการวิเคราะห์เชิงสถิติและเชิงพื้นที่ ซึ่งจะ

อธิบายต่อไป

4.1 ปริมาณข้อมูลทวิต

ในช่วงที่ทำการดึงข้อความทวิตในประเทศไทย พบว่าแต่ละ

วันระบบได้รับข้อความทวิตจำนวนกว่า 0.14 ล้านทวิต ทว่า

ด้วยขนาดฐานข้อมูลที่จำกัดและระยะเวลาที่ใช้ประมวลผลอาจ

มีสูงมากเกินไป ผู้วิจัยจึงจัดเก็บข้อมูลไว้ในระบบเพียง 15 วัน

ซึ่งรวมเป็นจำนวนประมาณ 2 ล้านข้อความ

4.2 การค้นหาข้อความ

เมื่อผู้ทดสอบกรอกคีย์เวิร์ดลงใน Text box ของหน้าเพจ และตั้งให้ค้นหา บราวเซอร์จะส่งคำร้องมายังเซิร์ฟเวอร์ด้วยพารามิเตอร์เหล่านี้

```
Keyword=<words>
bbox=<LAT_TL,LAT_BR,LON_BR,LON_TL>
startdate=<datetime> enddate=<datetime>
searchmode=< user or message>
```

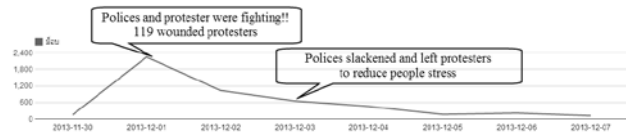
WOT จะทำการรวบรวมทวีตที่ตรงกับเงื่อนไขซึ่งแนบมากับคำร้องแล้วจัดข้อมูลทวีตให้อยู่ในรูปแบบ JSON (JavaScript Object Notation) แล้วจึงนำส่งคืนให้กับไคลเอนต์

4.3 การวิเคราะห์เชิงเวลา

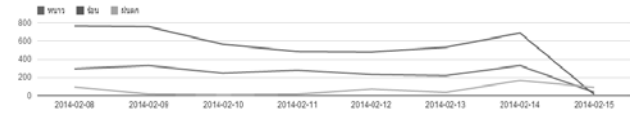
ผู้วิจัยได้ทำการทดลองค้นหาด้วยคีย์เวิร์ด “มือบ” “หนาว,ร้อน,ฝนตก” และ “รถติด” คีย์เวิร์ดแรกเกี่ยวกับเหตุการณ์ชุมนุมต่อต้านรัฐบาลตั้งแต่ปลายเดือนพฤศจิกายน ค.ศ. 2013 จะเห็นได้จากภาพที่ 4 ว่าในวันที่ 30 พ.ย. 2013 มีจำนวนทวีตที่มีคำว่า “มือบ” เพียง 149 ทวีต แต่ในวันที่ 1 ธ.ค. 2013 ซึ่งมีเหตุการณ์การปะทะกัน มีจำนวนถึง 2,246 ทวีต แล้วจึงลดลงในวันถัดๆ มา สำหรับคีย์เวิร์ด “หนาว,ร้อน,ฝนตก” เป็นการเปรียบเทียบความรู้สึกเกี่ยวกับอากาศของผู้ใช้ทวีตเตอร์ กราฟในภาพที่ 5 แสดงให้เห็นเส้นสีแทนคีย์เวิร์ดที่แตกต่างกันทำให้สามารถเปรียบเทียบข้อมูลระหว่างหัวข้อที่สนใจได้ง่ายขึ้น จะเห็นได้ว่ามีผู้ใช้คำว่า “ร้อน” มากกว่า “หนาว” และ “ฝนตก” ตลอดทั้งสัปดาห์ การค้นหาด้วยคำว่า “รถติด” โดยระบุช่วงเวลาทำให้เห็นความถี่ของทวีตที่เกิดขึ้นตามช่วงเวลาในหนึ่งวัน ว่าช่วงระหว่าง 9 โมงเช้าถึงเที่ยงมีการบ่นเรื่องรถติดมากกว่าช่วงเวลาอื่น ดังภาพที่ 6

4.4 การวิเคราะห์เชิงพื้นที่

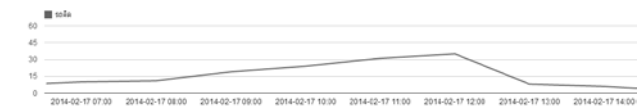
การแสดงตำแหน่งของทวีตบนแผนที่กูเกิลเอิร์ธทำให้เห็นการกระจายตำแหน่งของผู้ใช้ทวีตเตอร์ ทั้งยังเห็นคุณลักษณะที่ต่างกันในแต่ละหัวข้อ ตัวอย่างเช่น การค้นคำว่า “I'm at โรงเรียน” จะเห็นการกระจายของทวีตหนาแน่นในกรุงเทพฯ และปริมณฑล ซึ่งมีโรงเรียนจำนวนมาก และเบาบางในพื้นที่อื่นๆ เมื่อทำการขยายภาพเข้าไปจะเห็นว่าทวีตกระจุกตัวอยู่บนพื้นที่เดียวกันมากมายดังภาพที่ 7 ซึ่งเกิดจากผู้ใช้โพสต์ทวีตผ่าน



ภาพที่ 4: กราฟผลการค้นหาด้วยคีย์เวิร์ด “มือบ”



ภาพที่ 5: กราฟผลการค้นหาด้วยคีย์เวิร์ด “หนาว,ร้อน,ฝนตก”



ภาพที่ 6: กราฟผลการค้นหาด้วยคีย์เวิร์ด “รถติด”



ภาพที่ 7: ผลการค้นหาคำว่า “I’m at โรงเรียน” บนกูเกิลเอิร์ธ



ภาพที่ 8: ผลการค้นหาคำว่า “รถติด” บนกูเกิลเอิร์ธ

แอปพลิเคชันซึ่งใช้ชื่อสถานที่ในการระบุพิกัด ส่วนการค้นหาด้วยคำว่า “รถติด” จะเห็นหมุดเรียงเป็นแนวตามถนน ดังภาพที่ 8 เนื่องจากผู้ใช้โพสต์ข้อความอยู่บนถนนและได้พิกัดจาก GPS

นอกจากนี้ยังมีการคำนวณสถิติเชิงพื้นที่ เช่น ความเร็วในการแพร่กระจาย ระยะทางระหว่างทวีตที่เก่าที่สุดใหม่ที่สุดหรือไกลที่สุด ตามตารางที่ 1 ซึ่งเป็นผลจากการค้นหาคำว่า “รถติด” ในกรอบพื้นที่รอบประเทศไทย ซึ่งครอบคลุมกว่า 3,724,389.62 ตร.กม.ในช่วง 2 วันคือ 17-19 กุมภาพันธ์ 2014 ผู้ใช้งาน 350 คนทวีตข้อความที่เกี่ยวข้องถึง 913 ทวีตจากทั้งหมด 2,072,279 ทวีต โดยทวีตมีขอบเขตกว้าง 1,453.19

กม. และความเร็วในการแพร่กระจายไปจุดที่ไกลที่สุด 37.7 กม./ชม. หากนับจากทวีตที่เก่าที่สุดไปใหม่ที่สุดจะเห็นว่าความเร็วในการเดินทางคือ 11.64 กม./ชม.

ตารางที่ 1: ตัวอย่างตัวเลขผลลัพธ์จากการวิเคราะห์เชิงพื้นที่

Information	Keyword: รถติด
Total # tweets:	913 / 2072279
Total # twitter users:	350
Oldest tweet:	2014:2:17 00:00
Last tweet:	2014:2:19 23:49
Furthest of propagation (km):	1453.19
Speed of propagation (km/h):	37.77
Furthest of oldest tweet (km):	750.4
Speed from original (km/h):	11.64
Density estimation (tweets/km ²):	0.000245
Area (km ²):	3724389.62

5. บทสรุป

ผู้วิจัยได้พัฒนาระบบเว็บแอปพลิเคชันเพื่อศึกษาคุณลักษณะเชิงพื้นที่และเวลาจากข้อมูลทวีตเตอร์ ซึ่งพบรูปแบบหรือพฤติกรรมที่น่าสนใจและหลากหลาย และยังพบปัญหาบางอย่างที่ควรแก้ไขในอนาคต เช่น ทวีตเตอร์ไม่มีการแนบพิกัดสำหรับข้อความทวีต ทำให้รวบรวมข้อมูลได้ไม่ครอบคลุม

แนวทางในการพัฒนาต่อไปได้แก่ การวิเคราะห์ความหมายหรือความรู้สึกจากเนื้อหาในข้อความ การตรวจจับหัวข้อที่มีผู้โพสต์มากขึ้นอย่างรวดเร็ว การเคลื่อนที่ของฝูงชน (crowd mobility) เป็นต้น

เอกสารอ้างอิง

- [1] "Twitter," <http://en.wikipedia.org/wiki/Twitter> (March 2007)
- [2] Steven Millward, "Thailand Now Has 18 Million Social Media Users (INFOGRAPHIC)," <http://www.techinasia.com/thailand-18-million-social-media-users-in-2013> (May 2013)
- [3] J.M. Soler, F. Cuartero and M. Roblizo, "Twitter as a Tool for Predicting Elections Results," in Proceedings of IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, pp. 1194-1200, 2012.
- [4] S. Phuvipadawat and T. Murata, "Breaking News Detection and Tracking in Twitter," in Proceedings of IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, pp. 120-123, 2010.
- [5] J. Tuetrakul, "Social Network on Google Earth," Senior Project, Chulalongkorn University, 2011.
- [6] S. Doan, L. Ohno-Machado and N. Collier, "Enhancing Twitter Data Analysis with Simple Semantic Filtering: Example in Tracking Influenza-Like Illnesses," in Proceedings of IEEE Second Conference on Healthcare Informatics, Imaging and Systems Biology, 2012.
- [7] M. Demirbas, M.A. Bayir, C.G Akcora, Y.S. Yilmaz and H. Ferhatosmanoglu, "Crowd-Sourced Sensing and Collaboration Using Twitter," in Proceedings of World of Wireless Mobile and Multimedia Networks (WoWMoM), IEEE International Symposium, pp. 1-9, 2010.
- [8] A. Takeoka and U. Brajawidagda "Twitter Early Tsunami Warning System: A Case Study in Indonesia's Natural Disaster Management," in Proceedings of 46th Hawaii International Conference on System Sciences, pp. 2050-2060, 2013.
- [9] "Location, Location, Location" <https://blog.twitter.com/2009/location-location-location> (August, 2009)
- [10] A.A.A. Esmine, R.L. de Oliveira Jr. and S. Matwin "Hierarchical Classification Approach to Emotion Recognition in Twitter," in Proceedings of 11th International Conference on Machine Learning and Applications, pp. 381-385, 2012.
- [11] H. Gu, H. Hang, Q. Lv and D. Grunwald "Fusing Text and Friendships for Location Inference in Online Social Networks," in Proceedings of IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology, pp 158-165, 2012.
- [12] "Streaming API request parameters: Location" <https://dev.twitter.com/docs/streaming-apis/parameters#locations> (October, 2013)

เครือข่ายซับซ้อนและการจำแนกข้อมูลวิดีโอออนไลน์ Complex Networks and Online VDO Classification

สุวีวรรณ แจ้งจิต (Sureewan Jangjit)¹, นราธิป ทองปาน (Narathip Thongpan)²,

ศิริเรือง พัฒน์ช่วย (Siriruang Patchuay)³ และ นวพร วิสิฐพงษ์พันธ์ (Nawaporn Wisitpongphan)⁴

¹สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏกาญจนบุรี

²สาขาวิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏมหาสารคาม

³สาขาวิชาเทคโนโลยีวิศวกรรมคอมพิวเตอร์ คณะอุตสาหกรรมและเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลรัตนโกสินทร์

⁴สาขาวิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

itkan9@gmail.com, narathip.th@rmu.ac.th, siriruang.ph@hotmail.com, nawapornn@kmutnb.ac.th

บทคัดย่อ

การให้บริการในเครือข่ายสังคมออนไลน์ผ่านเว็บไซต์ เช่น การแบ่งปันรูปภาพ การดูวิดีโอออนไลน์ผ่านทางเฟสบุ๊คหรือ ยูทูปเป็นที่นิยมในปัจจุบัน ดังนั้นในงานวิจัยนี้จึงได้นำข้อมูลของผู้ใช้วิดีโอออนไลน์มาวิเคราะห์เพื่อ 1) หาความสัมพันธ์ของข้อมูลเพื่อนำเสนอในรูปแบบเครือข่ายซับซ้อนโดยใช้โปรแกรม Pajek และนำเสนอในรูปแบบภาพกราฟิกด้วยการจัดกลุ่มโดยใช้เทคนิค Self-Organizing Map (SOM) 2) วิเคราะห์หากฎความสัมพันธ์ของข้อมูลเพื่อนำไปประยุกต์ใช้ในการทำการตลาดเฉพาะกลุ่มโดยใช้เทคนิคต้นไม้ตัดสินใจ

ผลการวิจัยพบว่าผู้ใช้ที่มีระดับการศึกษาสูงและมีการจ้างงานแบบเต็มเวลามีการดูวิดีโอออนไลน์ผ่านทางสังคมออนไลน์มากที่สุดและต้นไม้การตัดสินใจเทคนิค ID3 ให้ค่าความถูกต้องในการทายผลระดับการศึกษายอดเยี่ยมซึ่งเท่ากับ 98.72 %

คำสำคัญ: วิดีโอออนไลน์ เครือข่ายสังคม การจัดกลุ่ม การจำแนกข้อมูล

Abstract

The social network services such as sharing photos and watching video online Facebook or YouTube are popular nowadays. In this study, we extracts video online user's information 1) present the relationship of user's characteristics in the form of complex network using Pajek program and presents a graphical user's classification using SOM technique 2) use decision tree techniques to identify decision rules that can be applied to the niche market research.

According to our findings, users with high level of education and who are, employed full time are those who watching video online via social media the most. In addition, the decision trees with ID3 technique yields to the results with highest accuracy - 98.72 %.

Keyword: video online, social network, clustering, classification.

1. บทนำ

การให้บริการในเครือข่ายสังคมออนไลน์ทำให้ผู้ใช้สามารถติดต่อสื่อสารกันได้อย่างไร้พรมแดนและหลากหลายรูปแบบ เช่น การแลกเปลี่ยนความคิดเห็น การแบ่งปันรูปภาพ วิดีโอ จากข้อมูลที่สร้างขึ้นเองหรือพบเจอจากสื่อต่าง ๆ ผ่านทางช่องทางออนไลน์โดยอาศัยอุปกรณ์มือถือหรือคอมพิวเตอร์ โดยกิจกรรมที่ได้รับความนิยมในสังคมออนไลน์ได้แก่การดูวิดีโอออนไลน์ ดังนั้นเพื่อค้นหาความสัมพันธ์ข้อมูลผู้ใช้งานวิดีโอออนไลน์ ในงานวิจัยนี้ได้นำเสนอเครือข่ายซับซ้อนของผู้ใช้งานวิดีโอออนไลน์ในรูปแบบกราฟที่เข้าใจง่าย เห็นกลุ่ม การเชื่อมโยงชัดเจนโดยใช้โปรแกรม Pajek และการจัดกลุ่มข้อมูลโดยวิธี SOM เพื่อแสดงการจัดกลุ่มด้วยภาพกราฟิกที่เข้าใจง่าย นอกจากนี้งานวิจัยนี้ยังได้นำเทคนิคต้นไม้ตัดสินใจ มาใช้ในการจำแนกข้อมูลของผู้ใช้งานวิดีโอออนไลน์เพื่อวิเคราะห์หาประเภทวิดีโอที่ได้รับความนิยมในแต่ละกลุ่มผู้ใช้งาน ซึ่งข้อมูลที่ได้จากงานวิจัยนี้สามารถนำมาใช้ในการทำการตลาดเฉพาะกลุ่มได้

2. วรรณกรรมที่เกี่ยวข้อง

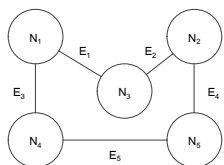
สื่อสังคมออนไลน์ (Social media) เป็นสังคมออนไลน์ที่ใช้สื่อ เช่น รูปภาพ ข้อความ วิดีโอและเสียงผ่านเครือข่ายสังคมออนไลน์ (Social network) เรียกว่า รูปแบบการมีปฏิสัมพันธ์ผ่านทางเครื่องมือบริการสื่อสังคมออนไลน์ (Social Network Service: SNS) เช่น เฟสบุ๊ก (Facebook) ทวิตเตอร์ (Twitter) เป็นต้น ที่สนับสนุนการใช้งานหลายกิจกรรม ได้แก่ การโพสต์ (Post) การอัปโหลด (Upload) การแบ่งปัน (Sharing) และการดู (Watch) โดยถูกออกแบบให้มีการใช้งานที่เข้าใจง่ายและไม่เสียค่าใช้จ่ายในการเข้าใช้งานส่งผลให้มีผู้ใช้เข้าไปทำกิจกรรมต่างๆ ในเครือข่ายสังคมออนไลน์เป็นจำนวนมากและมีข้อมูลเป็นจำนวนมากบนเครือข่ายดังกล่าวซึ่งสามารถนำไปใช้ในการวิจัยได้ โดยทฤษฎีและโปรแกรมที่สามารถนำมาประยุกต์ใช้เพื่อการประมวลผลข้อมูลบนเครือข่ายสังคมออนไลน์ มีดังนี้

2.1 โปรแกรม Pajek

โปรแกรม Pajek [1] เป็นโปรแกรมสำหรับวิเคราะห์และสร้างภาพ (Visualization) ในเครือข่ายขนาดใหญ่และซับซ้อน (Complex network) มีจุดเด่นในด้านการวิเคราะห์ชุดข้อมูลขนาดใหญ่ที่เป็นโปรแกรมซึ่งเปิดให้ใช้ฟรี (Free ware) มีพื้นฐานการแสดงผลและวิเคราะห์เครือข่ายซับซ้อนโดยอาศัยความรู้เรื่องกราฟ ซึ่งมีงานวิจัยหลายงานวิจัย อาทิงานวิจัยของ A. Khali et al. [2] ใช้โปรแกรมนี้ในการแสดงกราฟเพื่อให้เห็นการอ้างอิงทั้งหมดในรูปแบบต้นไม้รัศมี (radial tree form) และ C. Ni and Y. Ding [3] แสดงความสัมพันธ์ของรูปแบบข้อมูลของวารสาร LIS (Library and Information Science journals) เพื่อให้บรรณาธิการสามารถตรวจสอบการเชื่อมต่อของข้อมูลที่ใช้ประโยชน์ในกลุ่มวารสารใน LIS

2.2 ทฤษฎีกราฟ

กราฟ (Graph) ประกอบด้วยโหนด (Nodes) และเส้นเชื่อม (Edge) เขียนสัญลักษณ์ได้ $G = (N, E)$



ภาพที่ 1: แสดงส่วนประกอบของกราฟ

ภาพที่ 1 แสดงกราฟที่ประกอบด้วยโหนดจำนวน 5 โหนด เมื่อ $N(G) = \{N_1, N_2, N_3, N_4, N_5\}$, เส้นเชื่อมจำนวน 5 เส้น เมื่อ $E(G) = \{E_1, E_2, E_3, E_4, E_5\}$ โดยสามารถเขียนความสัมพันธ์ระหว่างโหนดคือ $E = \{\{N_1, N_3\}, \{N_2, N_3\}, \{N_1, N_4\}, \{N_2, N_5\}, \{N_4, N_5\}\}$

2.3 การทำเหมืองข้อมูล

เป็นการค้นหาความสัมพันธ์ที่ถูกซ่อนไว้ในข้อมูลจำนวนมากแบ่งการเรียนรู้เป็น 2 วิธีคือ วิธีที่ 1 การเรียนรู้แบบมีผู้สอน (Supervised Learning) และวิธีที่ 2 การเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) ดังนี้

วิธีที่ 1: การเรียนรู้แบบมีผู้สอนคือการกำหนดเป้าหมายที่ชัดเจน เพื่อทำนายค่าอินพุตที่เข้ามาใหม่ว่าเป้าหมายคืออะไร เรียกการเรียนรู้แบบนี้อาจเรียกข้อมูล (Classification) ซึ่งวิธีที่นิยมใช้เช่นเทคนิคต้นไม้ตัดสินใจ (Decision tree) เป็นเทคนิคที่นำเสนอข้อมูลโดยใช้โครงสร้างที่มีลักษณะคล้ายคลึงกับต้นไม้ โดยแต่ละโหนดแสดงถึงคุณลักษณะของข้อมูลที่ทดสอบ และที่โหนดใบ (Leaf node) บอกถึงคลาสที่จำแนกข้อมูลโดย R. Govindaraju and D. Mahardhika [4] ใช้การจำแนกข้อมูลเทคนิคต้นไม้ตัดสินใจ C4.5 เพราะเป็นวิธีที่ได้ถูกในการตัดสินใจ ตามที่ต้องการใช้งาน และเป็นวิธีที่แก้ไขได้ง่ายเพื่อนำมาใช้ในการพัฒนา web-based Decision Support System (DSS) ซึ่งประกอบด้วย learning module, application modules, exploration module

นอกจากนี้ยังมีการนำเทคนิค J48 หรือเรียกอีกชื่อว่า C4.5, เทคนิค JRip Rule และเทคนิค Neural Network มาวิเคราะห์รูปแบบการบุกรุกบนเครือข่าย Ji-roi-ni ภาคคีรีและกาญจนนา วิริยะพันธ์ [5] ซึ่งพบว่าเทคนิค J48 มีค่า 99.998% และ A. Cufoglu et al. [6] ได้แสดงการเปรียบเทียบประสิทธิภาพ Naïve Baysian, Instance-Based Learner, SimpleCART, NBTree, Id3, J48 และ SMO ในการจำแนกข้อมูล พบว่าวิธี NBTree ให้ผลการจำแนกที่ดีที่สุดคือ 90.20% และวิธีอื่นให้ผลที่ดีที่สุดคือเกิน 74% ที่ขนาดข้อมูล 1,000 ชุดและเปรียบเทียบการจำแนกข้อมูลที่มีขนาดข้อมูล 150, 327, 746 และ 1000 ชุดผลของการจำแนกข้อมูลมีค่าความถูกต้องสูงขึ้นเมื่อข้อมูลมีขนาดใหญ่ขึ้นตามลำดับ

วิธีที่ 2: การเรียนรู้แบบไม่มีผู้สอนคือไม่มีการกำหนดเป้าหมาย ไม่สามารถระบุจำนวนกลุ่มได้ว่ามีกี่กลุ่ม ต้องอาศัยการวัดความเหมือนหรือความแตกต่างของข้อมูล เรียกการเรียนรู้วิธีนี้ว่าการจัดกลุ่ม (Clustering) เช่น การจัดกลุ่มด้วยตัวเองของโคโฮเนน (Kohonen's Self-Organizing Map: SOM) โดย X. Rui and D. Wunsch II [7] ได้ศึกษาถึงคุณสมบัติการจัดกลุ่มให้เหมาะสมกับเงื่อนไขที่ต้องการ ควรคำนึงคุณสมบัติที่ส่งผลถึงประสิทธิภาพและประสิทธิผลในการจัดกลุ่ม เช่น รูปร่างของกลุ่มข้อมูล ข้อมูลขนาดใหญ่ มีมิติสูง ซับซ้อน ควรเลือกวิธีการจัดกลุ่มที่สนับสนุนกับเงื่อนไขของข้อมูล เป็นต้น

2.4 โปรแกรม Orange canvas

เป็น โปรแกรมด้านการทำเหมืองข้อมูล [8] ผ่านทางสัญลักษณ์ (Icon) ที่เป็นการสร้างแบบจำลองที่ครอบคลุมการทำงานด้านเหมืองข้อมูลซึ่งใช้งานง่ายและมีผลลัพธ์ที่แสดงในรูปแบบกราฟิกสวยงามและเป็นโปรแกรมฟรีสามารถนำไปใช้ได้กับทุกระบบปฏิบัติการโดยงานวิจัยนี้ใช้โปรแกรมนี้รุ่น 2.7

2.5 โปรแกรมเวก้า

โปรแกรมเวก้า (Waikato Environment for Knowledge Analysis: WEKA) มีข้อดีคือเป็นโปรแกรมฟรี สามารถทำงานได้กับทุกระบบปฏิบัติการ ง่ายต่อการใช้งานตลอดจนมีคำสั่งที่ครอบคลุมในด้านการทำเหมืองข้อมูล แต่โปรแกรมนี้มีจุดด้อยคือการแสดงผลกราฟิกไม่สวยงาม แต่โปรแกรมนี้ถือว่าเป็นที่นิยมใช้อย่างแพร่หลายเนื่องจากมีคำสั่งให้เลือกใช้อย่างครบถ้วนและสามารถเขียนฟังก์ชันเพิ่มในโปรแกรมได้โดยในงานวิจัยนี้ใช้โปรแกรมเวก้า รุ่น 3.6.9

ในโปรแกรมนี้การจำแนกข้อมูลด้วยต้นไม้ตัดสินใจมีเทคนิค การตัดกิ่งต้นไม้เพื่อช่วยลดความซับซ้อนของต้นไม้ และยังคงการจำแนกข้อมูลได้ถูกต้องโดยเทคนิค C4.5 วิธีการตัดกิ่งแบบ Reduced-error pruning: REP ที่ลดความผิดพลาดโดยใช้ชุดข้อมูลสำหรับตัดกิ่งแยกจากชุดข้อมูลฝึกสอนเพื่อใช้ตรวจสอบความถูกต้องของโหนดและใบของต้นไม้ที่ได้จากการสร้างต้นไม้ สามารถเลือกปรับค่าพารามิเตอร์ reducedErrorPruning เท่ากับ True และวิธีตัดกิ่งแบบ Error-based pruning: EBP โดยใส่ค่าความผิดพลาด วิธีนี้ใช้ชุดข้อมูลฝึกสอนในการสร้างและตัดกิ่งต้นไม้ตัดสินใจไม่แยกชุดสำหรับ

การตัดกิ่งสามารถเลือกปรับค่าพารามิเตอร์ reducedErrorPruning เท่ากับ False

2.6 การตลาดเฉพาะกลุ่ม

การตลาดเฉพาะกลุ่ม (Niche Market) เป็นการตลาดกลุ่มย่อยในตลาดใหญ่ เหมาะกับผู้ประกอบการที่มีทุนหรือทรัพยากร งบประมาณจำกัด มีการแข่งขันน้อยกว่าตลาดใหญ่ที่มุ่งเน้นคนจำนวนมาก (Mass market) ดังนั้นตลาดเฉพาะกลุ่มจึงมีการจำหน่ายสินค้าหรือให้บริการลูกค้าเฉพาะกลุ่มแบบเจาะจง โดยลูกค้ามีรสนิยมและความต้องการคล้ายคลึงกัน

การศึกษาการตลาดแบบเฉพาะกลุ่มอาทิ ระบบไร้สาย CATV, Video-on-demand [9] เช่น การประชุมทางไกลเป็นตลาดที่เกิดขึ้นใหม่และมีแนวโน้มและทิศทางที่สำคัญในอนาคต

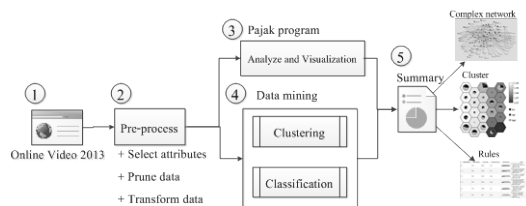
Pew Internet & American Life Project [10] เผยแพร่ข้อมูลผลการสำรวจกิจกรรมที่ผู้ใช้ในสังคมออนไลน์ทำบนสื่อออนไลน์ในเดือนกรกฎาคม 2013 บ่งชี้ว่าผู้ใช้ที่เป็นกลุ่มผู้ใหญ่ชาวอเมริกันมีกิจกรรมอัปโหลด (Upload) หรือโพสต์วิดีโอ (Post videos) เพิ่มขึ้นเป็นสองเท่าในระยะเวลา 4 ปี จาก 14% เป็น 31% ด้วยความนิยมที่เพิ่มขึ้น ผลลัพธ์แสดงสรุปอยู่ในรูปข้อความรูปแบบตาราง ทำให้ยากต่อการศึกษาเรื่องความสัมพันธ์ของแต่ละกิจกรรมในสังคมออนไลน์กับผู้ใช้งาน และมีการศึกษากิจกรรมการแบ่งปันวิดีโอ [11, 12] ที่ชี้ให้เห็นถึงความสัมพันธ์ของวิดีโอที่มีลักษณะเป็นเครือข่ายโลกใบเล็ก (small world networks) ซึ่งบ่งบอกถึงความสัมพันธ์ที่ใกล้เคียงกัน เนื่องจากความสอดคล้องกันระหว่างวิดีโอมีมาก (Strong correlation) ทำให้เป็นที่น่าสนใจในการศึกษาถึงข้อมูลในมุมมองอื่นที่เกี่ยวข้องกับเครือข่ายของวิดีโอ

ดังนั้นในงานวิจัยนี้จึงได้นำเสนอวิธีการแสดงความสัมพันธ์ของข้อมูล โดยใช้รูปแบบเครือข่ายซับซ้อนแสดงถึงความสัมพันธ์ของตัวแปรที่หลากหลาย ทำให้ได้ภาพที่น่าสนใจด้วยโปรแกรม Pajek นอกจากนี้ยังนำเสนอเทคนิคการจัดกลุ่มโดยใช้วิธี SOM ด้วยโปรแกรม Orange canvas ท้ายสุดจึงนำเสนอการจำแนกข้อมูลด้วยต้นไม้ตัดสินใจเพื่อให้ได้มาซึ่งกฎการตัดสินใจของการทำการตลาดเฉพาะกลุ่มของผู้ใช้เครือข่ายสังคมออนไลน์ โดยทำการเปรียบเทียบต้นไม้ตัดสินใจ

เทคนิค C4.5, ID3 และ NBTee เพื่อหาโมเดลที่ดีที่สุดและแม่นยำที่สุดในการจำแนกกลุ่มเพื่อทำการตลาด

3. ขั้นตอนวิธีการดำเนินการวิจัย

ในงานวิจัยนี้แบ่งการดำเนินงานเป็น 3 ส่วนคือ ส่วนที่ 1: แสดงรูปแบบกราฟความสัมพันธ์ของข้อมูลในรูปแบบเครือข่ายซับซ้อน ส่วนที่ 2: การจัดกลุ่มข้อมูลและส่วนที่ 3 การจำแนกข้อมูลซึ่งมีขั้นตอนการดำเนินงาน 5 ขั้นตอนดังภาพที่ 2



ภาพที่ 2: แสดงขั้นตอนการดำเนินงาน

ภาพที่ 2 ขั้นที่ 1 ศึกษาข้อมูลจาก [10] ซึ่งเป็นข้อมูลเกี่ยวกับผู้ใช้อินเทอร์เน็ตที่มีกิจกรรมวิดีโอออนไลน์ เช่น การดูวิดีโอออนไลน์ โดยข้อมูลมาจากการสัมภาษณ์ทางโทรศัพท์จากกลุ่มตัวแทนของผู้ใหญ่ชาวอเมริกันจำนวน 1,003 คน ที่อาศัยในสหรัฐอเมริกาโดย Princeton Survey Research Associates International (PSRAI) เป็นผู้เก็บข้อมูลระหว่าง 25-28 กรกฎาคม 2013

ขั้นที่ 2 เข้าสู่กระบวนการประมวลผลขั้นต้นเพื่อคัดเลือกแอททริบิวต์ที่ใช้ในการดำเนินงาน จากนั้นดำเนินการแปลงข้อมูลในแต่ละแอททริบิวต์ให้อยู่ในรูปแบบโปรแกรม Pajek และโปรแกรม Orange เพื่อใช้ประมวลผลประกอบด้วยแอททริบิวต์จำนวน 6 แอททริบิวต์ดังนี้

- 1) ช่องทางการดูวิดีโอได้แก่ การดูวิดีโอผ่านทางสังคมออนไลน์โดยใช้คอมพิวเตอร์ (use_FB) เช่น เฟสบุ๊ก และการดูวิดีโอผ่านโปรแกรมประยุกต์ของโทรศัพท์มือถือ (use_MB) เช่น Vine,
- 2) อายุ (age) ประกอบด้วย 4 ช่วงดังนี้ 18-29 ปี (1), 30-49 ปี (2), 50-64 ปี (3) และ 65 ปีขึ้นไป (4),
- 3) พื้นที่ที่ใช้สังคมออนไลน์ (region) ประกอบด้วย 4 พื้นที่ดังนี้ NORTH-EAST (1), MID-WEST (2), SOUTH (3) และ WEST (4),
- 4) เพศ (sex) ประกอบด้วยเพศชาย (1) และเพศหญิง (2),
- 5) การศึกษา (edu) ประกอบด้วย 3 ระดับคือ ระดับต่ำ (low) หมายถึงจบการศึกษาดำกว่ามัธยมศึกษาหรือเทียบเท่า

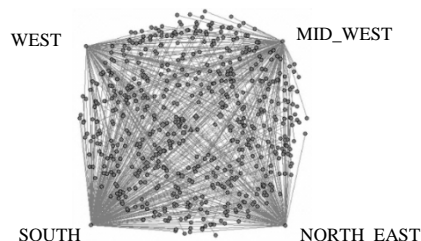
ระดับกลาง (mid) หมายถึงระดับอนุปริญญาหรือเทียบเท่า และระดับสูง (high) หมายถึงระดับปริญญาตรีขึ้นไป และ 6) ประเภทวิดีโอ (VOD01) และการจ้างงาน (employ) ประกอบด้วย 3 ระดับคือการจ้างแบบเต็มเวลา (Employed full-time) การจ้างบางส่วน (Employed part-time) และไม่มีการจ้าง (Not employed)

ขั้นตอนที่ 3 ใช้โปรแกรม Pajek วิเคราะห์ความสัมพันธ์และแสดงในรูปแบบกราฟ ขั้นที่ 4 ประกอบด้วย 2 ส่วนคือการจัดกลุ่มโดยใช้โปรแกรม Orange และจำแนกข้อมูลด้วยโปรแกรมเวก้า โดยเลือกชุดทดสอบเป็นชุดฝึกสอน (Use training set) โดยเป้าหมายการจำแนกข้อมูลด้วยต้นไม้ตัดสินใจคือต้องการผลลัพธ์ที่ใช้ทำนายผลระดับการศึกษาซึ่งจากงานวิจัยได้ศึกษาพบว่าเทคนิคต้นไม้ตัดสินใจที่ให้ประสิทธิภาพสูงในการจำแนกข้อมูลสูงและนิยมใช้คือเทคนิค C4.5, ID3 และ NBTee ดังนั้นในงานวิจัยนี้ได้เลือกใช้ 3 เทคนิคข้างต้นมาเปรียบเทียบและเลือกเทคนิคที่ให้ค่าความถูกต้อง (Accuracy) สูงสุดและมีค่าความคลาดเคลื่อนระหว่างค่าจริงกับค่าที่ใช้พยากรณ์ค่า (Root Mean Squared Error: RMSE) เนื่องจากค่าพยากรณ์มีค่าเข้าใกล้ค่าจริง และขั้นตอนสุดท้ายนำผลการวิเคราะห์มาสรุปผลลัพธ์ที่ได้จากการประมวลผลในขั้นที่ 3 และขั้นที่ 4

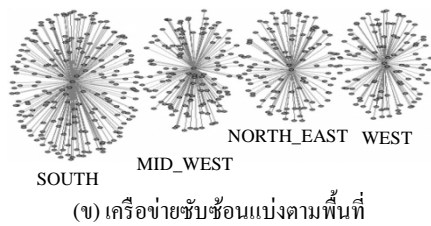
4. ผลการดำเนินงาน

ผลการดำเนินงานแบ่งเป็น 3 ส่วนได้แก่ ส่วนที่ 1 แสดงรูปแบบความสัมพันธ์แบบเครือข่าย ส่วนที่ 2 แสดงรูปแบบความสัมพันธ์ด้วยโปรแกรม SOM และส่วนที่ 3 จำแนกข้อมูลเพื่อทำนายผลระดับการศึกษา มีรายละเอียดดังนี้

ส่วนที่ 1: ผลการแสดงผลรูปแบบความสัมพันธ์แบบเครือข่ายด้วยโปรแกรม Pajek



(ก) เครือข่ายซับซ้อนแบ่งตามชนิดวิดีโอ ระดับการศึกษา ช่องทางการดูวิดีโอ การจ้างงาน และเพศ

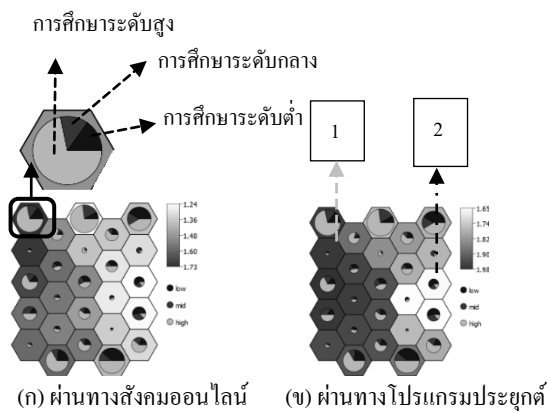


ภาพที่ 3: เครือข่ายซับซ้อนของผู้ใช้บริการสังคมออนไลน์

จากภาพที่ 3(ก) แสดงเครือข่ายซับซ้อนของข้อมูลระดับการศึกษา, ช่องทางการดูวิดีโอ, การจ้างงานและเพศโดยเห็นการเชื่อมโยงระหว่างกลุ่มข้อมูลชัดเจนและเป็นที่น่าสนใจ โดยพบว่าผู้ใช้ที่อยู่ในพื้นที่เดียวกันมีช่องทางการดูวิดีโอคล้ายกัน

ภาพที่ 3(ข) แสดงเครือข่ายซับซ้อนแบ่งตามพื้นที่ของผู้ใช้ โดยกลุ่มที่ดูวิดีโอออนไลน์มากที่สุดคือพื้นที่ SOUTH

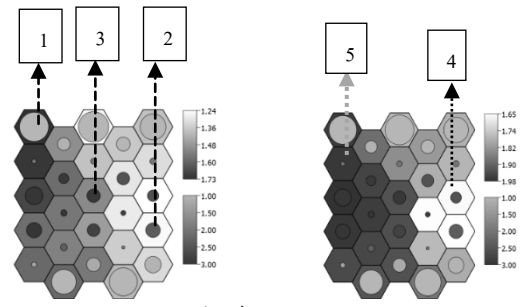
ส่วนที่ 2: ผลการแสดงผลรูปแบบความสัมพันธ์ด้วยโปรแกรม SOM ดังภาพที่ 4



ภาพที่ 4: ความสัมพันธ์ระหว่างช่องทางการดูวิดีโอ – ระดับการศึกษา

ภาพที่ 4 แสดงความสัมพันธ์ระหว่างช่องทางการดูวิดีโอและระดับการศึกษาโดยผู้ใช้ดูวิดีโอผ่านทางสังคมออนไลน์และผ่านทางโปรแกรมประยุกต์ของโทรศัพท์มือถือ โดยพื้นที่ของรูปหกเหลี่ยมอธิบายถึงช่องทางการดูวิดีโอผ่านทางสังคมออนไลน์หรือผ่านทางโปรแกรมประยุกต์ของโทรศัพท์มือถือ โดยหมายเลข 2 หมายถึงการดูวิดีโอออนไลน์และหมายเลข 1 หมายถึงไม่ดูวิดีโอออนไลน์

พบว่าผู้ใช้ที่มีการศึกษาระดับสูงเป็นกลุ่มที่ดูวิดีโอผ่านทางสังคมออนไลน์มากที่สุด โดยนิยมดูผ่านทางโปรแกรมประยุกต์ของโทรศัพท์มือถือและเมื่อต้องการศึกษาต่อว่าการจ้างงานมีความสัมพันธ์กับการดูวิดีโอออนไลน์หรือไม่ สามารถนำความสัมพันธ์ดังกล่าวมาแสดงได้ ดังภาพที่ 5



ภาพที่ 5: ความสัมพันธ์ระหว่างช่องทางการดูวิดีโอ – การจ้างงาน

จากภาพที่ 5 แสดงความสัมพันธ์ระหว่างช่องทางการดูวิดีโอและการจ้างงาน โดยผู้ใช้กระทำกิจกรรมผ่านทางสังคมออนไลน์และผ่านทางโปรแกรมประยุกต์ของโทรศัพท์มือถือโดยวงกลมภายในรูปหกเหลี่ยมอธิบายถึงการจ้างงานแทนด้วยหมายเลข 1 หมายถึงการจ้างงานแบบเต็มเวลา หมายเลข 2 หมายถึง การจ้างงานแบบบางส่วนและหมายเลข 3 หมายถึงการไม่มีงานทำ และพื้นที่ของรูปหกเหลี่ยมอธิบายถึงการดูวิดีโอผ่านทางสังคมออนไลน์หรือผ่านทางโปรแกรมประยุกต์ของโทรศัพท์มือถือ โดยหมายเลข 4 หมายถึงทำการดูวิดีโอออนไลน์ หมายเลข 5 หมายถึงการไม่ดูวิดีโอออนไลน์

พบว่าผู้ใช้ที่มีการจ้างงานแบบเต็มเวลาเป็นกลุ่มที่มีมากที่สุด โดยส่วนใหญ่ดูวิดีโอออนไลน์ผ่านทางสังคมออนไลน์มากกว่าผ่านทางโปรแกรมประยุกต์ของโทรศัพท์มือถือ

ผลลัพธ์จากภาพที่ 4 และภาพที่ 5 พบว่าผู้ใช้ที่มีระดับการศึกษาสูงและมีกรจ้างงานแบบเต็มเวลาเป็นกลุ่มที่มีมากที่สุด โดยดูวิดีโอออนไลน์ผ่านทางสังคมออนไลน์ เช่น เฟสบุ๊ก

ส่วนที่ 3: ในการจำแนกข้อมูลโดยเปรียบเทียบเทคนิคค้นไม่ตัดสินใจที่ 1

ตารางที่ 1: เปรียบเทียบผลลัพธ์เทคนิคค้นไม่ตัดสินใจ

Decision tree	C4.5 EBP	C4.5 REP	ID3	NBTree
Accuracy	70.88%	67.77%	98.72%	79.30%
RMSE	0.37	32.23	0.06	0.30
Build model (seconds)	0.07	0.01	0.01	3.25

ผลการทดสอบพบว่าเทคนิค ID3 ให้ค่าความถูกต้องสูงสุดคือ 98.72 % มีค่า RMSE น้อยสุดคือ 0.06 หมายความว่าผลของการทำนายมีค่าใกล้เคียงกับผลของค่าความเป็นจริงมากที่สุดโดยมีจำนวนโหนดใบที่เก็บคำตอบผลลัพธ์ของระดับการศึกษา

จำนวน 279 โหนด แบ่งเป็น โหนดใบเก็บค่าระดับการศึกษาสูง 123 โหนด, โหนดใบเก็บค่าระดับการศึกษากลาง 68 โหนดและ โหนดใบเก็บค่าระดับการศึกษาต่ำ 88 โหนด

จากต้นไม้ตัดสินใจที่ได้ ต้นไม้ที่มีกิ่งสั้น (shortest path) มีความสามารถจำแนกข้อมูลได้สูงกว่ากิ่งยาว (long path) ดังนั้น จึงทำการแปลงเป็นกฎจากต้นไม้ระดับ 0 (โหนดราก) ถึงระดับ 2 (โหนดใบมีค่าผลการทำนาย) ได้กฎจำนวน 12 กฎ ดังตารางที่ 2 เช่น ถ้าผู้ใช้อาศัยในพื้นที่ NORTH-EAST และทำงานแบบเต็มเวลาและเป็นเพศหญิงแล้วผลลัพธ์การทำนายคือมีระดับการศึกษาสูง เป็นต้น

ตารางที่ 2: แสดงกฎจากต้นไม้ตัดสินใจ 3 ระดับ

กฎที่ได้จากต้นไม้ตัดสินใจ 3 ระดับ				ผลลัพธ์
age_Range = 1	employ = 1	sex = 2	-	high
	employ = 2	cregion=4	Sports=1	mid
	employ = 3	cregion=2	Social_N=n	mid
age_Range = 2	Political=1	cregion=1	Animation=2	low
		cregion=1	Animation=1	high
age_Range = 3	employ=1	mobile_A=y	cregion=1	mid
		mobile_A=y	cregion=4	mid
	employ=2	Educational=1	Sports=2	high
		Educational=2	News=2	low
	employ=3	Music=1	Commercials=1	high
age_Range = 4	cregion=4	sex=1	-	high
		sex=2	News=2	high

5. สรุป

งานวิจัยนี้ได้ผลลัพธ์ดังนี้ 1) เครือข่ายซับซ้อนแสดงการเชื่อมโยงข้อมูลผู้ใช้สังคมออนไลน์มีลักษณะเป็นเครือข่ายโลกใบเล็ก กล่าวคือคนที่อยู่ในพื้นที่เดียวกันหรือใกล้กันมีโอกาสสนใจในกิจกรรมทางสังคมออนไลน์เหมือนกันมากกว่าคนที่อยู่พื้นที่ห่างไกล 2) การจัดกลุ่มพบว่าผู้ใช้ที่มีระดับการศึกษาสูงและมีการจ้างงานแบบเต็มเวลาเป็นกลุ่มที่มีมากที่สุดโดยดูวิถีออนไลน์ผ่านทางสังคมออนไลน์และ 3) เทคนิคต้นไม้ตัดสินใจ ID3 ให้ค่าการทำนายระดับการศึกษาถูกต้องสูงสุด 98.72% และเมื่อทำการแปลงต้นไม้ตัดสินใจ ID3 ที่ได้เป็นกฎการทำนายระดับการศึกษา เพื่อนำไปประยุกต์ใช้ในการทำตลาดเฉพาะกลุ่ม โดยใช้ผลการทำนายในการแบ่งกลุ่มลูกค้าเพื่อกำหนดกลุ่มเป้าหมายและนำข้อมูลที่ได้ไปประยุกต์ใช้การให้บริการ

หรือการออกแบบสินค้าเฉพาะกลุ่มเพื่อตอบสนองความต้องการของกลุ่มเป้าหมาย

งานวิจัยอนาคตผู้วิจัยจะจัดเก็บข้อมูลด้วยเทคนิคไยแมงมุม (Crawling) จากเว็บไซต์เกี่ยวกับวิดีโอออนไลน์เพื่อเพิ่มจำนวนข้อมูลให้มากขึ้นและเพิ่มประสิทธิภาพในการจัดกลุ่มและจำแนกข้อมูล

กิตติกรรมประกาศ

ขอขอบคุณแหล่งข้อมูลจาก Pew Internet & American Life Project [10] ที่อนุเคราะห์ข้อมูลในงานวิจัยนี้

เอกสารอ้างอิง

- [1] V. B. a. A. Mrvar. (2010). *Pajek Program for Analysis and Visualization of Large Networks*.
- [2] A. Khalid, et al., "Citation network visualization of CiteSeer dataset," in *Computer Sciences and Convergence Information Technology (ICCIT), 2011 6th International Conference on*, 2011, pp. 367-370.
- [3] C. Ni and Y. Ding, "Journal clustering through interlocking editorship information," presented at the Proceedings of the 73rd ASIS&T Annual Meeting on Navigating Streams in an Information Ecosystem - Volume 47, Pittsburgh, Pennsylvania, 2010.
- [4] R. Govindaraju and D. Mahardhika, "Web-based Decision Support System using C4.5 decision tree algorithm," in *Advanced Computer Science and Information System (ICACSIS), 2011 International Conference on*, 2011, pp. 395-400.
- [5] จิโรจน์ ภาคศิริและกาญจนา วิริยะพันธ์, "การวิเคราะห์รูปแบบการบุกรุกข้อมูลแบบเครือข่าย โดยใช้เทคนิคดาต้าไมน์นิ่ง," วารสารเทคโนโลยีสารสนเทศ, vol. ปีที่ 3 ฉบับที่ 6 กรกฎาคม - ธันวาคม 2550.
- [6] A. Cufoglu, et al., "A Comparative Study of Selected Classifiers with Classification Accuracy in User Profiling," in *Computer Science and Information Engineering, 2009 WRI World Congress on*, 2009, pp. 708-712.
- [7] X. Rui and D. Wunsch II, "Survey of clustering algorithms," *Neural Networks, IEEE Transactions on*, vol. 16, pp. 645-678, 2005.
- [8] J. Demšar, et al., "Orange: From Experimental Machine Learning to Interactive Data Mining," in *Knowledge Discovery in Databases: PKDD 2004*. vol. 3202, J.-F.

Boulicaut, *et al.*, Eds., ed: Springer Berlin Heidelberg, 2004, pp. 537-539.

- [9] H. H. Meinel, "The market for short-haul line-of-sight millimeterwave transmission links," in *Microwave Symposium Digest, 1996., IEEE MTT-S International*, 1996, pp. 487-489 vol.2.
- [10] K. Purcell. (2013). *Online Video 2013*. Available: <http://pewinternet.org/Reports/2013/Online-Video.aspx>
- [11] C. Xu, *et al.*, "Video sharing propagation in social networks: Measurement, modeling, and analysis," in *INFOCOM, 2013 Proceedings IEEE*, 2013, pp. 45-49.
- [12] C. Xu, *et al.*, "Statistics and Social Network of YouTube Videos," in *Quality of Service, 2008. IWQoS 2008. 16th International Workshop on*, 2008, pp. 229-238.

การรู้จำการเขียนตัวอักษรไทยด้วยลายมือบนโทรศัพท์เคลื่อนที่

Hand Written Thai Characters Recognition for Mobile Phone

ป๋จชัย พวงสุวรรณ (Patchai Phuangsuwan)¹ และ ชีระพงษ์ บริบริบูรณ์ (Theerapong Boriboon)²

^{1,2} สาขาวิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏอุดรธานี
patchai@udru.ac.th, pure_cs31@hotmail.com

บทคัดย่อ

วัตถุประสงค์ของงานวิจัยในครั้งนี้เพื่อพัฒนาโปรแกรมการรู้จำการเขียนตัวอักษรไทยด้วยลายมือบนโทรศัพท์เคลื่อนที่แบบระบบปฏิบัติการแอนดรอยด์ เป็นโปรแกรมเพิ่มเติมกับส่วนของแป้นพิมพ์ และเป็นอีกทางเลือกหนึ่งของผู้ใช้งานในการป้อนข้อมูล โดยใช้อัลกอริทึมโครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ (backpropagation neural network) ต้นไม้ตัดสินใจ (decision tree) แบบ C4.5 และซัพพอร์ตเวกเตอร์แมชชีน (support vector machine) โดยใช้คุณลักษณะเด่นของตัวอักษร ได้แก่ ตำแหน่งจุดเริ่มต้น จุดสิ้นสุดของตัวอักษร แบบของส่วนหัว การยก ขนาดความกว้างของตัวอักษร ลักษณะรอยหยักบนหัว บนตัวของตัวอักษร ลักษณะการม้วนของตัวอักษร และทิศทางของตัวอักษรจากรหัสโซ่ (chain code) การทดสอบความถูกต้องของแบบจำลองการรู้จำของข้อมูลทดสอบโดยใช้ 10 folds cross-validation ด้วยโปรแกรม Weka ความถูกต้องคิดเป็นร้อยละ 86.10, 87.57 และ 92.14 ตามลำดับ

คำสำคัญ: การรู้จำการเขียนตัวอักษรไทย เขียนตัวอักษรไทยด้วยลายมือ โทรศัพท์เคลื่อนที่ รหัสลูกโซ่ โครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ ต้นไม้ตัดสินใจ ซัพพอร์ตเวกเตอร์แมชชีน

Abstract

The purpose of this research is to develop the recognition program of handwritten Thai characters on Android mobile. This program is a supplement with the keyboard and works as an alternative user input. Recognition algorithms are back propagation neural networks, decision tree C4.5 and support vector machine by using features of the character, such as the beginning and end of the letter, header type, raising, width ratio, characteristic notches in the header and body of the letter, characteristics of the character rolls and orientation of

letter from chain code. The accuracy of recognition on testing data by using 10 folds cross-validation in Weka program are 86.10, 87.57 and 92.14% respectively.

Keyword: Thai characters recognition, hand-written, mobile phone, chain code, back-propagation neural network, decision tree, support vector machine

1. บทนำ

ปัญหาผู้ใช้งานบางคนที่การใช้โทรศัพท์เคลื่อนที่คือ การป้อนอักขระเข้าไปในโปรแกรมใช้งาน โดยเฉพาะอักขระภาษาไทยซึ่งมีจำนวนมาก และขนาดของปุ่มแป้นพิมพ์ที่มีขนาดเล็กทำให้ยากต่อการป้อนข้อมูลเข้าสู่โทรศัพท์ เนื่องจากว่าทุกคนคุ้นเคยการเขียนตัวอักษรซึ่งใช้ในชีวิตประจำวันอยู่แล้ว

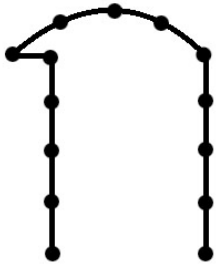
ดังนั้นทีมผู้วิจัยจึงมีแนวคิดที่จะพัฒนาโปรแกรมการรู้จำการเขียนตัวอักษรไทยด้วยลายมือบนโทรศัพท์เคลื่อนที่แบบระบบปฏิบัติการแอนดรอยด์ เพื่อเป็นโปรแกรมเพิ่มเติมกับส่วนของแป้นพิมพ์ และเป็นอีกทางเลือกหนึ่งของผู้ใช้งานในการป้อนข้อมูล

2. ทฤษฎีที่เกี่ยวข้อง

2.1 รหัสลูกโซ่ (Chain code)

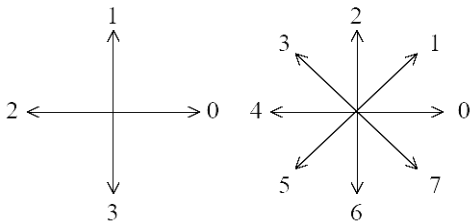
การอ่านจุดพิกเซลจากการเขียนเป็นวิธีหนึ่งที่ใช้ในการเก็บค่าจากการเขียนบนหน้าจอโทรศัพท์มือถือ โดยเมื่อกำหนดพิกเซลบนหน้าจอโทรศัพท์มือถือจะมีการเก็บจุดพิกเซลเริ่มต้นและเมื่อมีการลากนิ้วบนหน้าจอก็จะเก็บจุดพิกเซลที่นิ้วลากผ่านจนกระทั่งมีการยกนิ้วขึ้น โดยเก็บจุดพิกเซลสุดท้ายและนำจุดพิกเซลที่เก็บไว้มาหาทิศทางหรือมุมของตัวอักษรด้วยสมการต่อไปนี้ [1]

$$\theta = \arctan\left(\frac{y_2 - y_1}{x_2 - x_1}\right)$$



ภาพที่ 1: การอ่านจุดพิกเซลจากการเขียน ก-ไก่

Chain code [2] เป็นวิธีแสดงรูปร่างของวัตถุ สามารถนำไปใช้วิธีการอธิบายตัวอักษรในเทอมของลำดับของจุดที่ต่อเนื่องไปเป็นอักขระ โดยอาศัยการเปลี่ยนแปลงทิศทางของพิกเซลว่าจะจะไปไหนในทิศทางใด ซึ่งโดยทั่วไปจะเป็นแบบ 4 ทิศทาง หรือ 8 ทิศทางดังภาพที่ 2

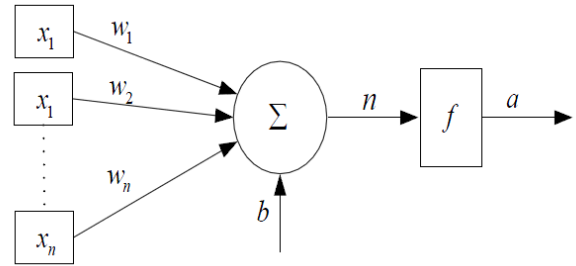


ภาพที่ 2: ทิศทางในการกำหนด Chain code

ตัวอย่างของตัวอักษรที่ถูกแปลงเป็น Chain code ของตัวอักษร “ก” แบบ 8 ทิศได้รหัส 2 2 2 2 2 4 1 1 7 7 6 6 6 6 6

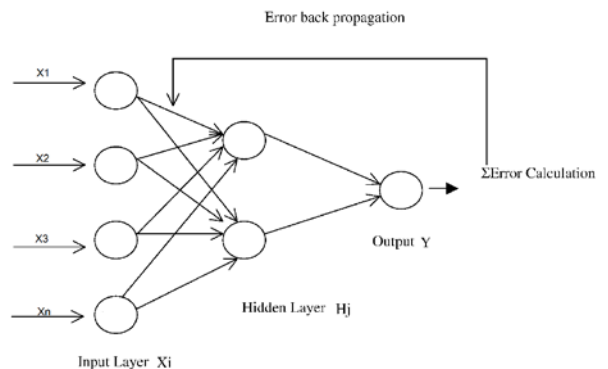
2.3 โครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ (Back propagation Neural Network)

โครงข่ายประสาทเทียม (Artificial Neural Network: ANN) มีพื้นฐานมาจากการจำลองการทำงานของสมองมนุษย์ด้วยโปรแกรมคอมพิวเตอร์[3-6] จุดมุ่งหมายของ โครงข่ายประสาทเทียมคือต้องการให้คอมพิวเตอร์มีความชาญฉลาดในการเรียนรู้เหมือนที่ มนุษย์มีการเรียนรู้ สามารถฝึกฝนได้ และสามารถนำความรู้และทักษะ รวมทั้งสามารถนำไปประยุกต์ใช้ได้ดีกับปัญหา Classification, Regression และ การเรียนรู้ของนิเวศเน็ตเวิร์ก ทำให้ได้โดยการส่งข้อมูลเข้ามายังส่วนที่เรียกว่าเพอร์เซ็ปตรอน (perceptron) สามารถเทียบได้กับเซลล์สมองของมนุษย์ โดยที่เพอร์เซ็ปตรอนทำการรับข้อมูลที่อยู่ในรูปของเวกเตอร์ของตัวเลข เข้ามาคำนวณ ดังภาพที่ 3



ภาพที่ 3: โครงข่ายประสาทเทียมเพอร์เซ็ปตรอน

โครงข่ายประสาทเทียมแบบแพร่หลายชั้น (Multilayer Perceptron:MLP) โครงข่ายประสาทเทียมแบบ MLP เป็นรูปแบบหนึ่งของโครงข่ายประสาทเทียมที่มีโครงสร้างเป็นแบบชั้นใช้สำหรับงานที่มีความซับซ้อนได้ผลเป็นอย่างดี โดยมีกระบวนการฝึกฝนเป็นแบบ supervise และใช้ขั้นตอนการส่งค่าย้อนกลับ (Back propagation) สำหรับการฝึกฝนกระบวนการส่งค่าย้อนกลับประกอบด้วย 2 ส่วนย่อยคือ การส่งผ่านไปข้างหน้า (Forward Pass) การส่งผ่านย้อนกลับ(Backward Pass) สำหรับการส่งผ่านไปข้างหน้า ข้อมูลจะผ่านเข้าโครงข่ายประสาทเทียมที่ชั้นของข้อมูลเข้าและจะส่งผ่านจากอีกชั้นหนึ่งไปสู่อีกชั้นหนึ่งจนกระทั่งถึงชั้นข้อมูลออก ส่วนการส่งผ่านย้อนกลับค่าน้ำหนักการเชื่อมต่อจะถูกปรับเปลี่ยนให้สอดคล้องกับกฎการแก้ข้อผิดพลาด (error correction) คือผลต่างของผลตอบที่แท้จริง (actual response) กับผลตอบเป้าหมาย (target response) เกิดเป็นสัญญาณผิดพลาด (error signal) ซึ่งสัญญาณผิดพลาดนี้จะถูกส่งย้อนกลับเข้าสู่โครงข่ายประสาทเทียมในทิศทางตรงกันข้ามกับการเชื่อมต่อ ค่าน้ำหนักการเชื่อมต่อจะถูกปรับและมีการเริ่มเรียนรู้ใหม่อีกครั้งจนกระทั่งผลตอบที่แท้จริงเข้าใกล้ผลตอบเป้าหมาย หรือเมื่อวนรอบอิมตัวแล้วกระบวนการเรียนรู้จะสิ้นสุดลง ดังภาพที่ 4



ภาพที่ 4: โครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ

$$y_j = f(x_1 w_{i,1}^n + x_2 w_{i,2}^n + x_3 w_{i,3}^n + \dots + x_i w_{i,j}^n + \theta_j^n)$$

$$y_j = f\left(\sum_{i=1}^n x_i w_{ij} + \theta_j\right)$$

โดยที่ y_j คือ ผลลัพธ์ในชั้นซ่อน หรือข้อมูลส่งออกในชั้นซ่อน โหนดที่ j w_{ij} คือ น้ำหนัก(weight) ของแต่ละเส้นเชื่อม (connection)

$f(u)$ คือ Activation function โดยใช้ sigmoid function มีสมการดังนี้

$$f(u) = \frac{1}{1 + e^{-u}}$$

x_i คือ ข้อมูลนำเข้าโหนดที่ i ในชั้นอินพุต

w_{ij} คือ น้ำหนักบนเส้นเชื่อมระหว่างโหนดที่ i ในชั้นอินพุตกับโหนด ที่ j ในชั้นซ่อน

θ_j คือ ค่าไบแอสของโหนดที่ j ในชั้นซ่อน

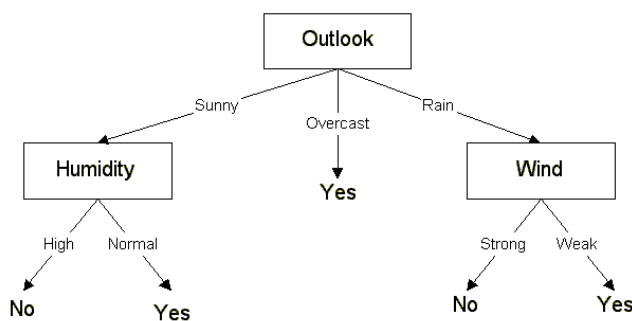
n คือ จำนวน โหนดทั้งหมดของชั้นอินพุต

e คือ ค่าคงตัวมีค่าประมาณ 2.71828

2.4 เทคนิคต้นไม้ตัดสินใจ (Decision Tree)

ต้นไม้ตัดสินใจเป็นเทคนิคหนึ่งของการจำแนกข้อมูล[6] ที่มีโครงสร้างคล้ายกับต้นไม้ ประกอบไปด้วย

- 1) โหนด (Node) เป็นส่วนของเงื่อนไขในการตัดสินใจ ซึ่งสามารถใช้ในการทำนายหรือพยากรณ์
- 2) Branch เป็นส่วนเชื่อมต่อระหว่างโหนดแสดงคุณลักษณะที่เป็นไปได้ทั้งหมดของแต่ละโหนด
- 3) Leaf Node เป็นส่วนที่แสดงกลุ่มที่ได้จากเงื่อนไขการตัดสินใจ



ภาพที่ 5: ตัวอย่างแบบจำลองต้นไม้ตัดสินใจ

การสร้างต้นไม้การตัดสินใจมีขั้นตอนในการสร้างประกอบด้วย 2 ขั้นตอน คือ

- 1) ขั้นตอนการสร้างต้นไม้เริ่มจากรากที่ประกอบด้วยจุดยอดที่มี Training Data ทั้งหมด และแยกกลุ่มตัวอย่างตามลักษณะประจำที่เลือกแล้วเรียกซ้ำ
- 2) ขั้นตอนการรวมและตัดแต่งกิ่ง จะรวมและกำจัดกิ่งที่มีผลกระทบต่อความถูกต้องน้อยออก

C4.5 เป็นอัลกอริทึมการจำแนกซึ่งให้ผลลัพธ์เป็นต้นไม้การตัดสินใจที่เป็นวิธีการสร้างต้นไม้เช่นเดียวกับ ID3 แต่เพิ่มกระบวนการปรับปรุงต้นไม้เพื่อให้มีลักษณะที่ดียิ่งขึ้น อัลกอริทึม C4.5 ถูกพัฒนาขึ้นมาโดย J. Ross Quinlan โดยใช้หลักการเลือกข้อมูลตามลำดับของค่า Information Gain สูงที่สุดเป็น Node เริ่มต้น และ Node ถัดไปจะมีค่า Information Gain ลดหลั่นกันไป สูตรที่ใช้ในการคำนวณหาค่า Information Gain ต้องเริ่มจากการหาค่า Entropy

Entropy :

$$\text{Entropy}(S) = \sum_{i=1}^c -p_i \log_2 p_i$$

โดย S คือ Attribute ที่นำมาวัดค่า Entropy

P_i คือ สัดส่วนของจำนวนสมาชิกของกลุ่ม i กับ

จำนวนสมาชิกทั้งหมดของกลุ่มตัวอย่าง

Information Gain :

$$\text{Gain}(S,A) = \text{Entropy}(S) - \sum_{v \in \text{Values}(A)} \frac{|S_v|}{|S|} \text{Entropy}(S_v)$$

โดย A คือ Attribute A

$|S_v|$ คือ สมาชิกของ Attribute A ที่มีค่า v

$|S|$ คือ จำนวนสมาชิกของกลุ่มตัวอย่าง

Split Info :

$$\text{SplitInfo}(T, A_i) = \sum_{k=1}^m (-1) \cdot p(T_{ij}) \cdot \log_2(p(T_{ij}))$$

Gain ratio :

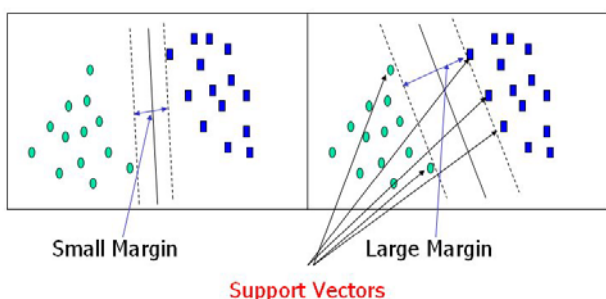
$$\text{Gain ratio} = \frac{\text{Gain}}{\text{SplitInfo}}$$

ปัญหา Over-Fitting เป็นปัญหาหนึ่งที่เกิดในการสร้างแบบจำลองต้นไม้ตัดสินใจ ลักษณะปัญหานี้คือแบบจำลองที่ได้มี Node และ Branch ที่มีความลึก ทำให้แบบจำลองซับซ้อนเกิน

ความจำเป็น หรือการที่จำนวนของข้อมูลฝึกสอนมีน้อยเกินกว่าที่จะจัดกลุ่มได้ตามต้องการ ซึ่งอัลกอริทึม C4.5 ได้ใช้เทคนิค Pruning เพื่อลดขนาดและความซับซ้อนของแบบจำลองต้นไม้ตัดสินใจ

2.5 ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine)

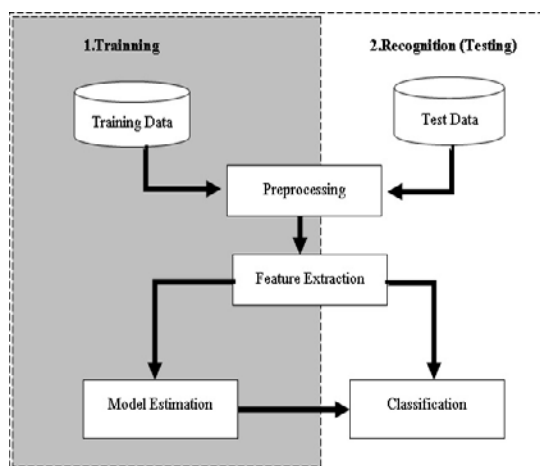
วิธีการของซัพพอร์ตเวกเตอร์แมชชีนจัดเป็นเทคนิคที่ใช้ในการแก้ปัญหาทางด้านการรู้จำรูปแบบ[8] โดยอาศัยหลักการของการหาสัมประสิทธิ์ของสมการเพื่อสร้างเส้นแบ่งแยกกลุ่มข้อมูลที่ถูกรูปแบบเข้าสู่กระบวนการสอนให้ระบบเรียนรู้ โดยเน้นไปยังเส้นแบ่งแยกแยะกลุ่มข้อมูลได้ดีที่สุด (Optimal Separating Hyperplane) ดังภาพที่ 6



ภาพที่ 6: ซัพพอร์ตเวกเตอร์แมชชีน

สำหรับรากฐานเดิมของซัพพอร์ตเวกเตอร์แมชชีน ถูกนำมาใช้กับข้อมูลที่เป็นเชิงเส้น แต่ในความเป็นจริงแล้วข้อมูลที่นำมาใช้ในระบบการสอนให้ระบบเรียนรู้ส่วนใหญ่มักเป็นข้อมูลแบบไม่เป็นเชิงเส้นซึ่งสามารถแก้ปัญหาดังกล่าวด้วยเคอร์เนลฟังก์ชัน (kernel function) ที่ใช้สำหรับการเปลี่ยนแปลงมิติของข้อมูลให้สูงขึ้น เพื่อช่วยในการเรียงตัวของข้อมูลเสียใหม่ที่เรียกว่า “ปริภูมิมิติสูง” (Higher Dimensional Space) เคอร์เนลที่นิยมใช้มี 3 ชนิด ด้วยกันคือ โพลีโนเมียล (Polynomial), เรเดียลเบสฟังก์ชัน (Radial Basis Function-RBF) และ ซิกมอยด์ฟังก์ชัน (Sigmoid) โดยงานวิจัยนี้ใช้เคอร์เนลโพลีโนเมียล

3. ขั้นตอนการดำเนินงาน



ภาพที่ 7: ขั้นตอนการดำเนินงานของระบบ

จากภาพ 7 อธิบายขั้นตอนการดำเนินงานของระบบได้ คือ จะแบ่งการทำงานออกเป็น 2 สิ่ง คือ 1) Training และ 2) Recognition (Testing) โดยในฝั่งของ Training จะมีข้อมูลสำหรับการฝึกฝน (Training Data) และฝั่งของ Testing จะเป็นฝั่งของการทดลองจากตัวแบบที่สร้างในขั้นตอน Classification

3.1 กระบวนการเตรียมการข้อมูล (Preprocessing)

กระบวนการเตรียมการข้อมูล ในกระบวนการนี้พิจารณาจากลักษณะการกหนดีวลงบนพื้นที่ของหน้าจอ โดยอาศัยเหตุการณ์ (Event) ที่เกิดขึ้นในการเก็บจุดพิกเซล จากนั้นนำจุดพิกเซลที่ได้มาหาตำแหน่งจุดเริ่มต้นจุดสุดท้าย การยก การหยัก หัวตัวอักษร การม้วน ลักษณะค่านวนหามุม และสร้างรหัส Chain code

3.2 นอร์มัลไลเซชัน (Normalization)

นอร์มัลไลเซชันเป็นการนำจุดพิกเซลที่ได้จากกระบวนการเตรียมข้อมูล (Preprocessing) ในการคัดเลือกจุดพิกเซลบางจุดในตำแหน่งต่างๆ ที่มีระยะห่างของจุดเท่ากัน มาคำนวณหารหัส Chain code แบบ 8 ทิศเพื่อแปลงรหัสเก็บในอาร์เรย์จำนวน 8 ช่อง เช่น Chaincode รหัส 0001122233444 เมื่อแปลงรหัสเก็บในอาร์เรย์จะได้ 32323000

3.3 การสกัดคุณลักษณะพิเศษ (Feature Extraction)

ลักษณะของตัวอักษรไทยสามารถนำมาพิจารณาหา ลักษณะพิเศษของตัวอักษร (Feature Extraction) ดังนี้ต่อไปนี้ ตำแหน่งจุดเริ่มต้น จุดสิ้นสุด หัวของตัวอักษร จำนวนการยก

อัตราส่วนความกว้างต่อความสูง ลักษณะรอยหยักบนส่วนหัวของตัวอักษร เช่น ฃ-ระฆัง ๗-มณโฑ ลักษณะรอยหยักของตัวอักษรลักษณะเช่น ต-เต่า ฒ-เต่า การม้วนของตัวอักษร ได้แก่ ม-ม้า ฃ-เนร ห-หีบ และรหัส Chain code ของตัวอักษร

3.4 การจำแนกข้อมูล (Classification)

การจำแนกหมวดหมู่ข้อมูล (Classification) ได้จากการนำข้อมูลที่ได้จากการเขียนบนหน้าจอโทรศัพท์มือถือระบบปฏิบัติการแอนดรอยด์มาแปลงเป็นไฟล์ ARFF ดังข้อมูลที่ใช้ตามกระบวนการหาคุณลักษณะพิเศษ (Feature Extraction) เพื่อนำมาสร้างตัวแบบตัวอักษรโดยใช้เทคนิคโครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ ด้วยฟังก์ชัน MultilayerPerceptron เทคนิคต้นไม้การตัดสินใจด้วยฟังก์ชัน Tree.J48 และเทคนิคซัพพอร์ตเวกเตอร์แมชชีนด้วยฟังก์ชัน LibSVM ในโปรแกรม Weka 3.7 เพื่อสร้างแบบจำลอง และทดสอบประสิทธิภาพของแบบจำลองตามข้อมูลข้อมูลสอน (Train) และข้อมูลทดสอบ (Test)

ข้อมูลที่ใช้ในการสร้างโมเดลมีข้อมูลเข้าจำนวน 10 แอดทริบิวต์ ตัวอักษรทั้งหมด 68 ตัวทั้งพยัญชนะ สระวรรณยุกต์ และสัญลักษณ์ จำนวนข้อมูลทั้งหมด 2640 รายการแบ่งเป็นข้อมูลที่ใช้ในการฝึกสอนและข้อมูลที่ใช้ในการทดสอบทำการทดสอบประสิทธิภาพโดยใช้ 10 folds cross-validation โดยมีการกำหนดพารามิเตอร์ของแต่ละเทคนิคดังนี้ 1) Back propagation กำหนดค่า hidden node=a (เท่ากับ 37) , learning rate =0.2 momentum=0.3 2) Tree.J48 confidence factor=0.2 3) Support vector machine กำหนดค่า c=0.1

4. ผลการดำเนินงาน

ผลจากการสร้างตัวแบบทั้ง 3 เทคนิคทั้งข้อมูลสอนและข้อมูลทดสอบได้ผลดังตารางที่ 1

ตารางที่ 1 ความถูกต้องของแต่ละเทคนิค (ร้อยละ)

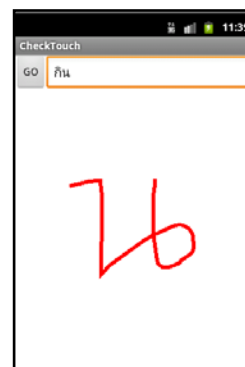
Technique	Train	Test
Back propagation	96.09	86.10
Decision tree	99.67	87.57
Support vector machine	99.26	92.14

จากตารางพบว่าประสิทธิภาพของตัวแบบที่ใช้เทคนิคซัพพอร์ตเวกเตอร์แมชชีนมีประสิทธิภาพมากที่สุด จึงได้นำค่า

Weight ที่ได้จากตัวแบบนำไปใช้ในโทรศัพท์มือถือโดยมีหน้าจอหลัก จะประกอบไปด้วยสองส่วนด้วยกัน ส่วนแรกเป็นช่องแสดงผลลัพธ์ที่ได้จากการเขียนกับปุ่มสำหรับส่งค่าไปยังโปรแกรมซึ่งจะอยู่ด้านบน ส่วนที่สองเป็นพื้นที่สีขาวสำหรับวาดตัวอักษร ตัวอย่างเช่น เมื่อมีการวาดตัวอักษร ก-ไก่ ตามด้วยสระอิ และ น-หนู จะแสดงผลลัพธ์ดังภาพที่ 7 และภาพที่ 8 ตามลำดับ



ภาพที่ 7: แสดงการเขียนตัวอักษร ก-ไก่



ภาพที่ 8: แสดงการเขียนตัวอักษรแบบเป็นคำ

5. บทสรุป

การวิจัยในครั้งนี้ใช้โครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ ต้นไม้การตัดสินใจและซัพพอร์ตเวกเตอร์แมชชีนในการรู้จำตัวอักษรไทยด้วยลายมือบนโทรศัพท์เคลื่อนที่แบบระบบปฏิบัติการแอนดรอยด์ ทำการตรวจสอบความถูกต้องโดยใช้ 10 folds cross-validation ผลปรากฏว่าความถูกต้องในการรู้จำของเทคนิคซัพพอร์ตเวกเตอร์แมชชีนดีที่สุด ร้อยละ 92.14 รองลงมาเป็นเทคนิคต้นไม้ตัดสินใจ ร้อยละ 87.57 และเทคนิคโครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ ร้อยละ 86.10 เนื่องจากตัวอักษรบางตัวมีลักษณะที่คล้ายคลึงกันทำให้การรู้จำมีความผิดพลาด ส่วนของการใช้งานโปรแกรมทั่วไปถือว่าอยู่ในเกณฑ์ดี

ใช้งานง่าย สามารถนำไปประยุกต์ใช้กับ โปรแกรมที่ขั้วบอร์คของ
มือถือระบบปฏิบัติการแอนดรอยด์ได้

เอกสารอ้างอิง

- [1] นฤมล ลับลิพ และศรุดิ อัสวเรืองสุข, “โปรแกรมอ่านภาพอักขระ
ภาษาไทยบน Android” วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต
มหาวิทยาลัยขอนแก่น.
- [2] Alasdair McAndrew, “Introduction to Digital Image
Processing with MATLAB”, pp.347-349, 2004.
- [3] Ravi Sheth, N C Chauhan, Mahesh M Goyani, Kinjal A
Mehta, “Handwritten Character Recognition System using
Chain code and Correlation Coefficient”, *IRCTITTCs 2011*,
pp.33-36, 2011.
- [4] สุรยุทธ ปรัชญา, “การรู้จำอักษรไทยโดยใช้โครงข่ายประสาทเทียม
แบบแพร่ย้อนกลับ” วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต
มหาวิทยาลัยเชียงใหม่.
- [5] Shraddha V. Shelke, D. M. Chandwadkar, “Automatic
System for Recognition of Handwritten Character using
Multiscale Neural Network”, *International Journal of
Computer Applications 2012*, pp. 89-92, 2012.
- [6] Brown, E.W., “Applying Neural Networks to
Character Recognition”, 1993. Available: [http://www.ccs.
neu.edu/home/feneric/charrecnn.html](http://www.ccs.neu.edu/home/feneric/charrecnn.html) (Accessed: 2013,
March, 13th).
- [7] J.R. Quinlan, “Introduction of Decision Trees”, *Machine
Learning*, vol. 1, pp.81-106, 1986.
- [8] Thanaruk Theeramunkong, “Introduction to Concepts and
Techniques in Data Mining and Application to Text
Mining”, pp.127-136, 2011.

การตามรอยเชิงฟีเจอร์สำหรับวิวัฒนาการสายผลิตภัณฑ์ซอฟต์แวร์ โดยใช้เวกเตอร์สเปซโมเดล

A Feature – Oriented Traceability for Software Product Line Evolution using Vector Space Model

กฤติกา ทิพย์ยอดศรี(Krittika Tipyodsree)¹ และ พรศิริ หมั่นไชยศรี (Pornsiri Muenchaisri)²

ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

¹Krittika.Ti@Student.chula.ac.th, ²Pornsiri.Mu@chula.ac.th

บทคัดย่อ

การตามรอยเชิงฟีเจอร์ในสายผลิตภัณฑ์ซอฟต์แวร์เป็นขั้นตอนที่ช่วยในการตรวจสอบความสอดคล้องและเชื่อมโยงกันระหว่างฟีเจอร์และส่วนประกอบของสิ่งประดิษฐ์ โดยจะมีลิงก์การตามรอยเชื่อมโยงกันระหว่างสิ่งประดิษฐ์เชิงซอฟต์แวร์ 2 สิ่งประดิษฐ์ เมื่อเกิดวิวัฒนาการสายผลิตภัณฑ์ซอฟต์แวร์ทำให้ลิงก์การตามรอยเดิมที่ถูกจัดเก็บไว้เปลี่ยนแปลงไป จึงทำให้ต้องมีการหาลิงก์การตามรอยใหม่ที่เกิดขึ้น ซึ่งวิธีเดิมจะทำการหาลิงก์การตามรอยใหม่โดยนักวิเคราะห์ระบบจะเป็นผู้วิเคราะห์และแก้ไขเอง ทำให้ใช้เวลานาน และเมื่อระบบมีขนาดใหญ่จะมีความซับซ้อนมากขึ้น จึงทำให้เกิดปัญหาการขาดการจัดการกับการเปลี่ยนแปลงของลิงก์การตามรอย ในงานวิจัยนี้จึงได้นำเสนอวิธีการตามรอยเชิงฟีเจอร์ในสายผลิตภัณฑ์ซอฟต์แวร์โดยใช้เทคนิคเวกเตอร์สเปซโมเดลเพื่อหาลิงก์การตามรอยใหม่ที่จะเกิดขึ้นแบบอัตโนมัติ เพื่อช่วยให้นักวิเคราะห์ระบบและผู้พัฒนาสามารถที่จะดูแลจัดการชิ้นส่วนซอฟต์แวร์ในระบบได้อย่างมีประสิทธิภาพ ลดค่าใช้จ่าย ประหยัดเวลา จากผลการทดลองที่ได้พบว่าวิธีการดังกล่าวให้ผลลัพธ์ที่ถูกต้อง ครบถ้วนและใช้เวลาน้อยกว่าวิธีที่ผ่านมา ซึ่งทำการประเมินผลโดยใช้ค่าความแม่นยำและค่าความระลึก

คำสำคัญ: เวกเตอร์สเปซโมเดล,สายผลิตภัณฑ์ซอฟต์แวร์,การตามรอย

Abstract

The features-oriented traceability in software product line are a process that helps to ensure consistency and coherence between features and components of the

architecture and have a link to trace links between software artifacts two architectures. Software product line evolution made exit links to trace changes. We need to find a new link to the trace. The old method without an automated tool,a software analyst have to find links to a new trace by analyzed and fix it that take longer. When the system is large, it is more complicated. The problem is lack of deal with the dynamics of link to the trace. This research proposes a way to trace the features in a software product line using the vector space model to find links to the new trace automatically.An approach help systems analysis and developers to manage the software components in the system effectively, reduce costs, save time. The results shows that such an approach results accurately and in less time than past methods. Evaluated using the precision and the recall.

Keywords: Vector Space Model, Software Product Line, Traceability

1. คำนำ

สายผลิตภัณฑ์ซอฟต์แวร์เป็นแนวคิดที่เน้นการนำกลับมาใช้ใหม่(Reuse)โดยจะต้องสกัดหาส่วนที่มีความเป็นทั่วไปของผลิตภัณฑ์(Commonality)และความจำเพาะหรือความผันแปรของผลิตภัณฑ์(Variability) เพื่อหาสินทรัพย์หลัก (Core Asset)ของกลุ่มของผลิตภัณฑ์ในสายผลิตภัณฑ์ซอฟต์แวร์ การพัฒนาระบบสายผลิตภัณฑ์ซอฟต์แวร์ระบบใดระบบหนึ่งขึ้นส่วนต่างๆของซอฟต์แวร์ในแต่ละเฟสของการพัฒนาย่อมมี

ความสอดคล้องและเชื่อมโยงกันและสามารถตรวจสอบย้อนกลับได้คือการตามรอยนั่นเองและจะมีลิงก์การตามรอยระหว่างชิ้นส่วนต่างๆที่เชื่อมโยงกันอยู่ ยกตัวอย่างในเฟสของการวิเคราะห์และออกแบบ เช่น แผนภาพยูสเคสจะมีลิงก์การตามรอยไปยังแผนภาพกิจกรรม เพื่อที่จะสามารถสร้างแผนภาพกิจกรรมได้แผนภาพกิจกรรมจะมีลิงก์การตามรอยไปยังแผนภาพคลาสเพื่อที่จะสร้างแผนภาพคลาสได้และแผนภาพต่างๆสามารถตรวจสอบกลับไปยังข้อกำหนดของความต้องการหรือพีเจอร์ของระบบได้ จะเห็นว่าทุกๆชิ้นส่วนของระบบต้องมีความสอดคล้องและไม่ขัดแย้งกันจึงจะสามารถส่งต่อให้เฟสของการพัฒนาและการทดสอบสามารถพัฒนาระบบออกมาอย่างมีประสิทธิภาพและสอดคล้องกับความต้องการของผู้ใช้งานอย่างสูงสุด

ดังนั้นการเปลี่ยนแปลงความต้องการของผู้ใช้ภายหลังจากการพัฒนาระบบ หรือภายหลังวิวัฒนาการของสายผลิตภัณฑ์ย่อมส่งผลให้ลิงก์ระหว่างชิ้นส่วนต่างๆที่ถูกจัดเก็บไว้อย่างถูกต้องและสอดคล้องกันมีการเปลี่ยนแปลงด้วย การขาดการจัดการกับการเปลี่ยนแปลงของลิงก์การตามรอยระหว่างพีเจอร์และส่วนประกอบของสิ่งประดิษฐ์ อันเนื่องมาจากวิวัฒนาการของสายผลิตภัณฑ์ทำให้เกิดการเปลี่ยนแปลงภายในแบบจำลองพีเจอร์และส่วนประกอบของสิ่งประดิษฐ์ และเกิดการเปลี่ยนแปลงระหว่างลิงก์การตามรอยที่เชื่อมโยงกันระหว่างพีเจอร์และส่วนประกอบของสิ่งประดิษฐ์ด้วย และส่งผลให้การตามรอยระหว่างพีเจอร์และส่วนประกอบของสิ่งประดิษฐ์ที่ถูกจัดเก็บไว้ไม่ถูกต้องและไม่สอดคล้องกับความต้องการที่แท้จริงของผู้ใช้

งานวิจัยนี้จึงได้นำเสนอเทคนิคเวคเตอร์สเปซโมเดล ซึ่งเป็นวิธีที่ช่วยค้นหาลิงก์การตามรอยใหม่ที่อาจเกิดขึ้นระหว่างพีเจอร์และส่วนประกอบของสิ่งประดิษฐ์ภายหลังการเกิดวิวัฒนาการของสายผลิตภัณฑ์

สำหรับรายละเอียดในส่วนต่อไปในหัวข้อที่ 2 และ 3 จะกล่าวถึงงานวิจัยและทฤษฎีที่เกี่ยวข้องตามลำดับ ในหัวข้อที่ 4 จะกล่าวถึง วิธีการดำเนินงานวิจัย หัวข้อที่ 5 กล่าวถึงการทดลองและการประเมินผล และหัวข้อสุดท้ายคือหัวข้อที่ 6 เป็นการสรุปผลงานวิจัยและแนวทางดำเนินการในอนาคต

2. งานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ทำการศึกษางานวิจัยที่เกี่ยวข้องกับการตามรอยสายผลิตภัณฑ์ซอฟต์แวร์โดยมีวัตถุประสงค์เพื่อเป็นแนวทางในการทำงานวิจัยให้มีประสิทธิภาพให้เกิดผลลัพธ์ที่มีความสมบูรณ์

Xiaofan Chen และ John Grundy [9]งานวิจัยเสนอวิธีการในการตามรอยลิงก์การเปลี่ยนแปลงในสายผลิตภัณฑ์ซอฟต์แวร์โดยอัตโนมัติระหว่างเอกสารและโค้ดใช้เทคนิคการค้นคืน Vector Space Model ในการตามรอยโดยผลลัพธ์ที่ได้จะได้ลิงก์การตามรอยระหว่างเอกสารและโค้ด โดยในงานวิจัยนี้ได้ปรับปรุงประสิทธิภาพของ Vector Space Model ให้มีความถูกต้องและความแม่นยำมากขึ้น โดยใช้วิธีการรวมกันของ 3 เทคนิคเข้ามาช่วยแก้ปัญหา คือ Regular Expression Key Phrases และการจัดกลุ่มผลการทดลองแสดงให้เห็นว่าวิธีการของผู้วิจัยแสดงให้เห็นว่าช่วยเพิ่มประสิทธิภาพการทำงานของ Vector Space Model ดีขึ้น

Kamthorn Chavanaves [10] งานวิจัยเสนอวิธีฟังก์ชันลิงก์การตามรอยที่เป็นไปได้เพื่อใช้ในการค้นหาลิงก์การตามรอยที่สามารถเกิดขึ้นใหม่ระหว่างพีเจอร์กับส่วนประกอบของสิ่งประดิษฐ์เชิงซอฟต์แวร์ภายหลังการเกิดวิวัฒนาการของสายผลิตภัณฑ์ซอฟต์แวร์ โดยใช้เทคนิคกิ่งของผลิตภัณฑ์และแบบจำลองความผันแปรเชิงตั้งฉากมาใช้โดยจะทำการวิเคราะห์ว่าถ้าพีเจอร์ที่มีคุณลักษณะเป็นจุดที่เกิดความแปรปรวนของกลุ่มของพีเจอร์กลุ่มใดกลุ่มหนึ่งในแบบจำลองพีเจอร์ มีการตามรอยไปยังส่วนประกอบของสิ่งประดิษฐ์ที่มีคุณลักษณะเป็นจุดที่เกิดความแปรปรวนของกลุ่มของส่วนประกอบของสิ่งประดิษฐ์กลุ่มใดกลุ่มหนึ่งในสิ่งประดิษฐ์เชิงซอฟต์แวร์พีเจอร์ที่มีคุณลักษณะเป็นตัวแปรของกลุ่มของพีเจอร์ที่มีการตามรอยอยู่นั้นจะมีการตามรอยไปยังส่วนประกอบของสิ่งประดิษฐ์เชิงซอฟต์แวร์ที่มีคุณลักษณะเป็นตัวแปรของกลุ่มของส่วนประกอบของสิ่งประดิษฐ์ที่มีการตามรอยอยู่ด้วย

Sayphin Taweessri [11]งานวิจัยเสนอวิธีการสร้างการติดตามความต้องการโดยใช้การคำนวณพจน์ความคล้ายกันซึ่งเป็นการสนับสนุนการติดตามความต้องการหรือการติดตามความสัมพันธ์ระหว่างชิ้นส่วนของซอฟต์แวร์ในระยะเวลาของการวิเคราะห์ความต้องการและการออกแบบซอฟต์แวร์โดย

อัตโนมัติใช้เทคนิคการค้นคืน vector space model เพื่อคำนวณค่าความคล้ายกันสำหรับการสร้างการติดตามความต้องการ

3. ทฤษฎีที่เกี่ยวข้อง

3.1 สายผลิตภัณฑ์ซอฟต์แวร์ (Software Product Line)[1]

เป็นการนำแนวคิดการผลิตเชิงอุตสาหกรรมมาใช้กับซอฟต์แวร์ เพื่อให้ Software สามารถ reuse ได้ง่ายโดยนำหลักการที่สำคัญ คือ การจัดกับสายการผลิต จะเริ่มต้นตั้งแต่การเก็บรวบรวมความต้องการ วิเคราะห์และออกแบบผลิตภัณฑ์วางแผนสายการผลิต จนถึงเริ่มต้นผลิตสินค้า

3.2 การตามรอย (Traceability)[3]

คือการเชื่อมโยงหรือระบุความสัมพันธ์ระหว่างสิ่งประดิษฐ์เชิงซอฟต์แวร์ 2 สิ่งประดิษฐ์ จะช่วยให้เราสามารถตรวจสอบได้ว่าซอฟต์แวร์ที่เราพัฒนาขึ้นมาได้ถูกพัฒนาตามความต้องการของผู้ใช้อย่างแท้จริงหรือไม่

3.3 การค้นคืนสารสนเทศ (Information Retrieval)[4]

การค้นคืนสารสนเทศคือการรวบรวม การจัดเก็บ และการเข้าถึงข้อมูลโดยพิจารณาและวิเคราะห์ถึงความสัมพันธ์การขึ้นต่อกัน ผลกระทบหรือความคล้ายกันของเอกสาร เพื่อรวบรวมและจัดเก็บให้ตรงกับความต้องการหรือคิวรี (Queries) มีการสร้างดัชนี (Indexing) ในการเข้าถึง

3.4 เวกเตอร์สเปซโมเดล (Vector Space Model)[4]

เป็นโมเดลในการจัดเก็บเอกสาร โดยจะแทนเอกสารที่ผู้ใช้คิวรีเข้ามาด้วยเวกเตอร์จากนั้นทำการเปรียบเทียบโดยทำการคำนวณความคล้ายกันเพื่อหาเอกสารที่มีเวกเตอร์คล้ายกับของผู้ใช้มากที่สุดออกมาเป็นผลลัพธ์ ในการนับความถี่ของคำที่ปรากฏและการให้ค่าน้ำหนักของคำสำคัญในเอกสารดังสมการที่ 1 [4]

$$wi = tfi * \log(D/dfi) \quad (1)$$

เมื่อ wi แทนค่าน้ำหนักของคำสำคัญ i ในเอกสาร

tfi แทนความถี่ หรือ จำนวนคำสำคัญที่ปรากฏในเอกสาร
 dfi แทนผลรวมค่าความถี่ของคำสำคัญ i ในทุกๆเอกสารที่อยู่ในฐานข้อมูล

D แทนจำนวนเอกสารที่อยู่ในฐานข้อมูล

และคำนวณค่าความคล้ายกัน (Similarity) ของเอกสารดังสมการที่ 2

$$S = ((dXd') / (|d||d'|)) \quad (2)$$

เมื่อ d แทนเวกเตอร์ของเอกสาร, d' แทนเวกเตอร์ของคิวรี,
 $|d|$ แทนขนาดของเอกสาร, $|d'|$ แทนขนาดของคิวรี

สามารถวัดประสิทธิภาพของการค้นคืนสารสนเทศวัดได้จากค่าความแม่นยำ (Precision) และความระลึก (Recall)

P = (จำนวนเอกสารที่เกี่ยวข้องเนื่องที่ค้นคืนได้/จำนวนเอกสารทั้งหมดที่ทำการค้นคืนออกมาได้) (3)

R = (จำนวนเอกสารที่เกี่ยวข้องเนื่องที่ค้นคืนได้/จำนวนเอกสารทั้งหมดที่เกี่ยวข้องเนื่องทั้งหมดในฐานข้อมูล) (4)

3.5 การจัดการความผันแปร (Variability Management) [5]

ความผันแปรในสายผลิตภัณฑ์ซอฟต์แวร์แบ่งออกเป็น 3 ประเภท ได้แก่ 1. คอมมอนอลิตี (Commonality) เป็นลักษณะใดๆที่แสดงคุณลักษณะที่มีในทุกผลิตภัณฑ์ในสายผลิตภัณฑ์ซอฟต์แวร์ 2. แวริบิลิตี (Variability) เป็นลักษณะใดๆที่แสดงคุณลักษณะที่มีในบางผลิตภัณฑ์ในสายผลิตภัณฑ์ซอฟต์แวร์แต่ไม่ทั้งหมด 3. โปรดักต์สเปซิฟิค (Product-Specific) เป็นลักษณะใดๆที่แสดงคุณลักษณะที่มีเฉพาะในผลิตภัณฑ์หนึ่งๆ

3.6 วิธีการวิเคราะห์โดเมนเชิงฟีเจอร์ (Feature-Oriented Domain Analysis (FODA) Method) [6]

ฟีเจอร์คือลักษณะใดๆที่แสดงคุณลักษณะของระบบ ฟีเจอร์ในแบบจำลองฟีเจอร์สามารถแบ่งออกได้เป็น 3 ประเภท เพื่อที่จะระบุความผันแปรของฟีเจอรันั้นว่าเป็นความผันแปรประเภทใด ฟีเจอร์จำเป็น (Common Feature) ฟีเจอร์ทางเลือก (Optional Feature) ฟีเจอร์เลือกอย่างใดอย่างหนึ่ง (Alternative Feature)

3.7 วิศวกรรมสายผลิตภัณฑ์บนพื้นฐานยูเอ็มแอลหรือพลัส (Product Line UML-Based Software Engineering) [7]

ในการสร้างแบบจำลองสายผลิตภัณฑ์ด้วยยูเอ็มแอลจะมีส่วนที่แตกต่างจากยูเอ็มแอลทั่วไปเล็กน้อยได้แก่

1. แผนภาพยูสเคสเป็นแผนภาพที่นำเสนอให้เห็นการปฏิสัมพันธ์หรือการติดต่อระหว่างผู้ใช้งานกับระบบ โดยยูสเคสของสายผลิตภัณฑ์ซอฟต์แวร์จะมีการกำหนดสเตอริโอไทป์ (Stereo Type) เพื่อบอกประเภทของยูสเคสว่าเป็นแบบทั่วไป (Common Use Case) แบบเลือกได้ (Optional Use Case) แบบทางเลือก (Alternative Use Case)

2.แผนภาพคลาส เป็นแผนภาพยูเอ็มแอลที่นำเสนอการทำแบบจำลองของระบบเชิงวัตถุที่จะมีการกำหนดคสเตอร์ไอโทป์ (Stereo Type)เพื่อบอกประเภทของคลาส 3 แบบ ประกอบด้วย คลาสแบบทั่วไป(Common Class) คลาสแบบเลือกได้ (Optional class) และคลาสแบบทางเลือก (Alternative class)

3.แผนภาพลำดับ เป็นแผนภาพที่แสดงปฏิสัมพันธ์ ระหว่างกันของวัตถุตามลำดับของเหตุการณ์ที่เกิดขึ้น ณ เวลาต่างๆ โดยจะแสดงถึงลำดับของการส่งสาร (Message)

4. การตามรอยเชิงพีเจอร์สำหรับวิวัฒนาการสายผลิตภัณฑ์ซอฟต์แวร์โดยใช้เทคนิคเวกเตอร์สเปซโมเดล

วิธีการสร้างลิงก์การตามรอยใหม่แบ่งออกเป็น 6 ขั้นตอนดังภาพที่1 โดยข้อมูลนำเข้าของวิธีการตามรอยนี้คือ แผนภาพยูสเคส แผนภาพคลาสแผนภาพลำดับ พีเจอร์ในแบบจำลองพีเจอร์ ข้อมูลส่งออกของวิธีการตามรอยนี้คือ ผลของข้อมูลการตามรอยระหว่างแบบจำลองพีเจอร์และสิ่งประดิษฐ์เชิงซอฟต์แวร์



ภาพที่1: ภาพรวมของระบบ

- 1.การสร้างข้อมูลการตามรอยของพีเจอร์ โดยจัดเก็บอยู่ในตารางพร้อมทั้งมีหมายเลขกำกับและกำหนดคสเตอร์ไอโทป์
- 2.สร้างข้อมูลการตามรอยของส่วนประกอบของสิ่งประดิษฐ์

ได้แก่ แผนภาพยูสเคส แผนภาพคลาส และแผนภาพซีเควนซ์ และทำการจัดเก็บคุณลักษณะในตารางพร้อมทั้งมีหมายเลขกำกับและกำหนดคสเตอร์ไอโทป์

3.สร้างลิงก์การตามรอยระหว่างพีเจอร์และส่วนประกอบของสิ่งประดิษฐ์ โดยผู้พัฒนาระบบจะทำการสร้างลิงก์การตามรอยระหว่างพีเจอร์และส่วนประกอบของสิ่งประดิษฐ์

4.ทำการเพิ่ม ลด แก้ไขพีเจอร์ เมื่อมีพีเจอร์ใหม่ถูกเพิ่มเข้ามา หรือมีการแก้ไข ลบพีเจอร์จะทำโดยผู้พัฒนาระบบ

5.ทำการเพิ่ม ลด แก้ไขส่วนประกอบของสิ่งประดิษฐ์ จะต้องมีการจัดการส่วนประกอบของสิ่งประดิษฐ์ที่ถูกเพิ่มเข้าไปใหม่ ที่ถูกแก้ไขและลบออก

เมื่อมีการเปรียบเทียบทั้งสองเวอร์ชันของ โปรแกรมแล้วพบว่ามีการเปลี่ยนแปลงเกิดขึ้น ทั้งการเพิ่ม ลดและแก้ไข ข้อมูลทั้งหมดนี้จะถูกเก็บไว้ในตาราง จะใช้ค่าที่อยู่ในพีเจอร์ V.2 มาใช้เป็นค่าค้น ในการสอบถามและสร้างลิงก์ในขั้นตอนถัดไปแบบอัตโนมัติ

ตารางที่ 1: ตัวอย่างการเปรียบเทียบพีเจอร์ทั้งสองเวอร์ชัน

พีเจอร์ V.1	พีเจอร์ที่ถูกแก้ไข	ประเภทการแก้ไข	พีเจอร์ V.2
Transmission	Transmission	-	Transmission
Horsepower	Null	Delete	-
Air Conditioning	Air Conditioning	-	Air Conditioning
Manual	Manual	-	Manual
Automatic	Self-driving	Edit	Self-driving
Null	Cool	Add	Cool
Null	Warm	Add	Warm

6.สร้างลิงก์การตามรอยใหม่ที่อาจเกิดขึ้นภายหลังวิวัฒนาการของสายผลิตภัณฑ์จะใช้สูตรเวกเตอร์สเปซโมเดล ในการช่วยสร้างลิงก์การตามรอยใหม่ที่เป็นไปได้ที่อาจเกิดขึ้นระหว่างพีเจอร์และส่วนประกอบของสิ่งประดิษฐ์ภายหลังวิวัฒนาการของสายผลิตภัณฑ์โดยการค้นหาชื่อที่มีความคล้ายกัน ซึ่งเริ่มต้นด้วยการวิเคราะห์คำสำคัญและตัดคำที่ไม่พิจารณาเพื่อสร้างกรณีการค้นคืน สร้างตารางกรณีสำหรับจัดเก็บคำสำคัญ จากนั้นหาความถี่และน้ำหนักคำสำคัญจากสมการ(จากหัวข้อ3.4)

$$wi = tfi * \log(D/dfi) \tag{1}$$

จากตัวอย่างกำหนดให้มีสิ่งประดิษฐ์ที่ทำการจัดเก็บทั้งหมด มี 3 สิ่งประดิษฐ์ (D=3) และพีเจอร์หรือคิวรีที่ต้องการค้นคือ Q: “Transmission fluid high” ดังตารางที่ 2

ตารางที่ 2: ตัวอย่างการหาความถี่และน้ำหนักคำสำคัญ

Term Vector Model Based on $w_i = t_{fi} * IDF_i$											
Query, Q: “Transmission fluid high”											
D1 : Transmission fluid is high											
D2 : <i>Airconditioning</i> are the process of altering the property of air											
D3 : A <i>manual transmission</i> , also known as a gearbox											
D = 3 , $IDF = \log\left(\frac{D}{df_i}\right)$											
		Counts, t_{fi}						Weights, $w_i = t_{fi} * IDF_i$			
Term	Q	D	D	D	d_i	$\frac{D}{df_i}$	IDF_i	Q	D1	D2	D3
		1	2	3							
Air	0	0	1	0	1	3	0.5	0	0	0.5	0
AirConditioning	0	0	1	0	1	3	0.5	0	0	0.5	0
Altering	0	0	1	0	1	3	0.5	0	0	0.5	0
fluid	1	1	0	0	1	3	0.5	0.5	0.5	0	0
high	1	1	0	0	1	3	0.5	0.5	0.5	0	0
Gearbox	0	0	0	1	1	3	0.5	0	0	0	0.5
known	0	0	0	1	1	3	0.5	0	0	0	0.5
manual	0	0	0	1	1	3	0.5	0	0	0	0.5
process	0	0	1	0	1	3	0.5	0	0	0.5	0
property	0	0	1	0	1	3	0.5	0	0	0.5	0
Transmission	1	1	0	1	2	1.5	0.2	0.5	0.5	0	0.5

คำนวณค่าระดับความคล้ายกันระหว่างชิ้นส่วนซอฟต์แวร์ดังสมการ(จากหัวข้อ3.4)

$$S = ((dXd')/(|d||d'|)) \tag{2}$$

ตารางที่ 3: ค่าระดับความคล้ายกัน

ชิ้นส่วนของซอฟต์แวร์ (D)	ชิ้นส่วนของซอฟต์แวร์ (D')	ค่าระดับความคล้ายกัน	เปอร์เซ็นต์ค่าระดับความคล้ายกัน
Q	D1	1	100%
Q	D2	0	0%
Q	D3	0.2887	28.87%

ในการสร้างลิงค์ใหม่จะพิจารณาจากค่าเปอร์เซ็นต์ระดับความคล้ายกันของชิ้นส่วนซอฟต์แวร์ที่จับคู่กันมีค่าสูงที่สุด หลังจากนั้นจะมีการกระทำซ้ำทุกข้อมูลที่มีการจัดเก็บไว้จนได้ลิงค์ใหม่ทั้งหมดที่จะเกิดขึ้นระหว่างพีเจอร์โมเดลและสิ่งประดิษฐ์เชิงซอฟต์แวร์ภายหลังวิวัฒนาการของสายผลิตภัณฑ์

5. การทดลองและการประเมินผล

การทดลองทำโดยนำผลิตภัณฑ์ซอฟต์แวร์ระบบรถยนต์และระบบเอทีเอ็มมาทำการประเมินผลโดยจะจัดเก็บข้อมูลการตามรอย หลังจากนั้นทำการเพิ่ม ลด แก้ไขพีเจอร์ ระบบจะทำการค้นหาลิงค์การตามรอยที่เกิดขึ้นใหม่โดยอัตโนมัติโดยเปรียบเทียบระหว่างผลการตามรอยทั้งที่เกิดขึ้นจริงกับเทคนิคเวกเตอร์สเปซโมเดล ดังตารางที่ 4 และทำการประเมินด้วย ค่าความระลึกละและค่าความแม่นยำ พบว่าค่าของ ค่าความระลึกละและค่าความแม่นยำที่เกิดขึ้นมีค่าเท่ากับ 1 คิดเป็นร้อยละ 100 ซึ่งมีความแม่นยำสูง ดังตารางที่ 5

ตารางที่ 4: ตัวอย่างลิงก์การตามรอยที่เกิดขึ้นหลังวิวัฒนาการของสายผลิตภัณฑ์ด้วยวิธีเวกเตอร์สเปซโมเดลซอฟต์แวร์ระบบรถยนต์และระบบเอทีเอ็ม

ระบบ	Feature ID	Feature Name	Class ID	Class Name	Stereo Type
Car	43	Transmission	61	Transmission	Common
	46	Air Conditioning	65	Air Conditioning	Optional
	47	Manual	66	Manual	Alternative
	51	Warm	73	Warm	Alternative
Atm	25	Transaction	33	Transaction	Common
	40	Account	56	Account	Common
	53	Pin	16	Pin	Common
	35	Amount	54	Amount	Alternative

ตารางที่ 5: ตัวอย่างการประเมินผลการตามรอยโดยใช้เทคนิคเวกเตอร์สเปซโมเดล

ระบบ	คำสำคัญ/ชั้นส่วนของซอฟต์แวร์	เอกสารทั้งหมดที่ค้นคืนได้ (1)	เอกสารที่ถูกต้องที่ค้นคืนได้ (2)	เอกสารที่ถูกต้องทั้งหมดในฐานข้อมูล (3)	ค่าเฉลี่ย (2)/(3)	ค่าความแม่นยำ (2)/(1)
Car	Transmission	8	8	8	1.00	1.00
	Air Conditioning	2	2	2	1.00	1.00
	Manual	10	10	10	1.00	1.00
	Warm	5	5	6	0.83	1.00
Atm	Transaction	15	15	15	1.00	1.00
	Account	19	19	19	1.00	1.00
	Pin	5	5	5	1.00	1.00
	Amount	11	11	11	1.00	1.00
					7.83	8.00
Average					1.00	1.00

6. สรุปผลงานวิจัยและแนวทางดำเนินการในอนาคต

งานวิจัยนี้สามารถสร้างลิงก์การตามรอยใหม่ที่เป็นไปได้ที่อาจเกิดขึ้นระหว่างฟีเจอร์และส่วนประกอบของสิ่งประดิษฐ์ภายหลังวิวัฒนาการของสายผลิตภัณฑ์โดยมีค่าความแม่นยำและค่าความระลึกรมีความถูกต้องครบถ้วนยกเว้นในกรณีค่าผสมที่ประกอบด้วยค่าหลายๆค่าและค่าที่มีรากศัพท์เดียวกันแต่เขียนแตกต่างกัน ยังไม่สามารถค้นหาโดยใช้วิธีนี้ได้

แนวทางดำเนินงานต่อไปคือการปรับปรุงการค้นหาสิ่งใหม่ที่เกิดขึ้น โดยสามารถนำวิธีเวกเตอร์สเปซโมเดลไปปรับปรุงให้มีประสิทธิภาพขึ้นเพื่อให้สามารถค้นหาสิ่งที่เกิดขึ้นได้ครบถ้วนครอบคลุมมากยิ่งขึ้น

เอกสารอ้างอิง

- [1] Clement P, Northop L., "Software Product Lines: Practices and Patterns" ,Boston, MA: Addison Wesley Longman Inc., 2001.
- [2] Eriksson, Magnus. "An approach to software product line use case modeling". (2006).
- [3] Devis, A.M., "Software Requirement: Objects, Functions and States." ,Upper Saddle River, NJ, USA: Prentice-Hall, 1993.
- [4] Gerard Salton and Michale J. McGill. "Introduction to Modern Information Retrieval", Auckland: McGraw-Hill International Book, 1983.
- [5] Schmid, K., van der Linden, F., and Rommes, E., "Software Product Lines in Action: The Best Industrial Practice in Product Line Engineering. Puducherry", India: Integra Software Services Pvt., 2007.
- [6] Kang, K., Cohen, S., Hess, J., Novak, W., and Peterson, "A Feature-Oriented Domain Analysis (FODA) Feasibility Study". Software Engineering Institute, Technical report, CMU/SEI-90-TR-021, 1990.
- [7] Hassan Gomaa, "Designing Software Product Lines with UML", Addison-Wesley, 2004
- [8] E. Harold, "XML Bible", IDG Books Worldwide, An International Data Group Company, 1999.
- [9] Xiaofan Chen ,John Grundy, "Improving Automated Documentation to Code Traceability by Combining Retrieval Techniques", ASE, 2011. IEEE/ICCSA, 2007

- [10] Kamthon Chavanaves ,“A Feature-Oriented Traceability for Software Product Line Evolution”,*IEEE/ICCSIT*,2011
- [11] สายพิน ทวีศรี,“การสร้างการติดตามความต้องการโดยใช้การคำนวณพจน์ความคล้ายกัน”, *ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย*,2006

การจัดเก็บข้อมูลไบนารีในฐานข้อมูลในรูปแบบ Base128

Storing Binary Data to Database using Base128 encoding

จรรยาภรณ์ ประสมสิทธิ์ (Chanyaphon Prasomsat)¹ และชัยพร เขมะภาคะพันธ์ (Chaiyapon Khemapatapan)²

สาขาวิศวกรรมคอมพิวเตอร์และโทรคมนาคม คณะวิศวกรรมศาสตร์ มหาวิทยาลัยธุรกิจบัณฑิตย์
wasana.nui@gmail.com¹, chaiyaporn@dpu.ac.th²

บทคัดย่อ

งานวิจัยนี้เป็นการพัฒนาวิธีการจัดเก็บข้อมูลไบนารี ให้มีประสิทธิภาพมากขึ้น ซึ่งพัฒนาบนเว็บเบราว์เซอร์แอปพลิเคชันกับระบบฐานข้อมูลโดยวิธีที่พัฒนามีการใช้พื้นที่ในการจัดเก็บข้อมูลลดลงด้วยการแปลงเป็นข้อมูลแบบ Base128 ที่อยู่ในรูปของสตริง ซึ่งจะช่วยลดความเสี่ยงที่อาจทำให้ฐานข้อมูลเสียหาย และง่ายต่อการค้นหาข้อมูลในฐานข้อมูล โดยการใช้เทคนิคการแปลงข้อมูลไบนารีให้เป็นข้อมูลสตริง ด้วยอินเด็กซ์ของข้อมูล การทดสอบจะทำการเปรียบเทียบระหว่างการจัดเก็บข้อมูลในรูปแบบ BLOB Base32 Base64 และ Base128 ตามที่เสนอ

ผลการทดสอบพบว่า การจัดเก็บข้อมูลแบบ Base128 ใช้พื้นที่การจัดเก็บข้อมูลใกล้เคียงกับการจัดเก็บข้อมูลแบบ BLOB ในขณะที่ใช้เวลาในการแปลงและบันทึกมากกว่าการจัดเก็บข้อมูลแบบ Base64 เล็กน้อย ส่วนการค้นหาข้อมูลและแปลงกลับใช้เวลาใกล้เคียงกับการจัดเก็บข้อมูลแบบ Base64

คำสำคัญ: Base128 Base64 BLOB Database

Abstract

This research is to develop an efficiency method of storing binary data which is developed on Web browser application with database system. The proposed method aims to reduce storage size by converting binary data into Base128 data. Storing Base128 string data in database will reduce the risk of database damage and help to find information in database. The conversion will use a binary value to index Base128 string. The testing results between various encoding: BLOB, Base32, Base64 and Base128 are compared.

The results show that storing data using Base128 consumes storage size nearby storing data using BLOB. However, Base128 encoding will consume time to convert and store more slightly than Base64 encoding. Fortunately, Base128 encoding uses time to query and revert into binary data close to Base64 encoding.

Keyword: Base128, Base64, BLOB, Database

1. บทนำ

ปัจจุบันการใช้งานคอมพิวเตอร์และการใช้บริการอินเทอร์เน็ตมีจำนวนมากขึ้นอย่างรวดเร็ว ทำให้ข้อมูลที่อยู่ในระบบมีมากขึ้น ส่งผลให้การบันทึกหรือจัดเก็บข้อมูลจากสื่อต่างๆ ต้องใช้ปริมาณความจุข้อมูลเพิ่มขึ้นตามไปด้วย แต่การจะเพิ่มหรือพัฒนาอุปกรณ์บันทึกข้อมูลให้มีขนาดความจุสูงขึ้นนั้นทำได้ไม่ง่ายนัก เนื่องจากต้องใช้การลงทุนสูง และการค้นคว้าวิจัยด้านอุปกรณ์การจัดเก็บข้อมูลนั้นกระทำได้อย่างยากลำบากกับข้อมูลในฐานข้อมูลเป็นสิ่งที่สำคัญมาก หากเกิดความเสียหายหรือข้อผิดพลาดใดๆ ขึ้น จะทำให้ส่งผลเสียใหญ่หลวงต่อระบบได้ ซึ่งการจัดเก็บข้อมูลในปริมาณมากนั้น สามารถส่งผลทำให้เกิดความเสียหายต่อฐานข้อมูลนั้นๆ ได้

ที่ผ่านมา ข้อมูลในฐานข้อมูลส่วนใหญ่ สามารถแยกได้เป็นข้อมูลประเภทไฟล์ซึ่งอยู่ในรูปแบบของข้อมูลไบนารี ข้อมูลประเภทข้อความ และข้อมูลตัวเลข จะเห็นว่าปริมาณข้อมูลไบนารีนั้น เป็นข้อมูลที่มีความยุ่งยากในการจัดเก็บข้อมูลมากที่สุด และสามารถส่งผลให้ฐานข้อมูลเกิดความเสียหายต่อการจัดเก็บในระบบฐานข้อมูลได้ โดยปกติสามารถจัดเก็บข้อมูลไบนารีในรูปแบบของการเก็บ path ของไฟล์ข้อมูล และทำการจัดเก็บข้อมูลจริงไว้ในรูปแบบของไฟล์แทน นอกจากนี้ก็สามารถจัดเก็บในรูปแบบของ BLOB ก็ได้ ซึ่งช่วยให้การจัดการกับ

ข้อมูลเป็นแบบเบ็ดเสร็จในระบบฐานข้อมูล อย่างไรก็ตามการ จัดเก็บข้อมูลแบบ BLOB ยังมีความเสี่ยงของความเสียหาย ของระบบฐานข้อมูลได้ เนื่องจากการจัดเก็บข้อมูลไบนารี ลงฐานข้อมูลโดยตรง ซึ่งอาจมีรหัส ASCII [1] ที่ทำให้ระบบ ฐานข้อมูลทำงานผิดพลาดได้ แต่ถ้าเทียบกับการเก็บข้อมูลแบบ พาท (path) แล้วจะมีข้อเสีย คือ การบริหารจัดการข้อมูลเพิ่มขึ้น เนื่องจากข้อมูลไม่ได้เก็บเป็นก้อนเดียวกัน

โดยงานวิจัยนี้ได้พัฒนาในด้านการแปลงข้อมูลไบนารีให้ อยู่ในรูปแบบสตริง โดยผ่านกระบวนการเข้ารหัสข้อมูลแบบ ไบนารีแปลงเป็นข้อมูลประเภทสตริง โดยใช้วิธีการสร้างจาวา คลาสขึ้นมาเรียกว่า Base128 เพื่อลดขนาดของข้อมูลไบนารีให้ ได้ใกล้เคียงขนาดจริงให้มากที่สุด โดยมีแนวทางการวิจัยที่ พัฒนามาจากการเข้ารหัสข้อมูลแบบ Base64 [2],[3] ซึ่งเป็น มาตรฐานของการเข้ารหัสข้อมูลไบนารีเป็นข้อมูลสตริงที่ใช้ งานกันทั่วไป [4-6] แล้ว

2. งานวิจัยที่เกี่ยวข้อง

2.1 การบันทึกข้อมูลไบนารี

งานวิจัยในนี้ มีวัตถุประสงค์ที่จะเข้ารหัสข้อมูลชนิดไฟล์ให้ อยู่ในไฟล์แบบสตริงก่อนเก็บลงฐานข้อมูล โดยมุ่งเน้นการลด ขนาดของข้อมูลไบนารี เมื่อบันทึกข้อมูลลงฐานข้อมูลให้ได้ ขนาดใกล้เคียงกับขนาดข้อมูลจริงให้มากที่สุด เมื่อเปรียบเทียบกับ การเข้ารหัสข้อมูลแบบ Base32 [2] และ Base64 ซึ่งก่อให้เกิดผลคือ ขนาดของข้อมูลไบนารีมีขนาดเล็กลง การจัดเก็บข้อมูลมีความเร็วและยืดหยุ่นขึ้น และสามารถค้นหา ข้อมูลไบนารีต่างๆ โดยค้นหาจากข้อความที่เป็นเนื้อหาของไฟล์ ข้อมูลนั้นๆ แทนการค้นหาภาพโดยใช้อินเด็กซ์ (Index) ซึ่งทำให้มีความสะดวกและรวดเร็วยิ่งขึ้น ซึ่งเทคนิคที่ใช้ในงานวิจัยนี้ คือ การแทนข้อมูลไบนารีที่เป็นชนิดไบต์ (Byte) เป็นชนิด สตริงมีข้อดีคือ การจัดเก็บข้อมูลมีความเร็วในการจัดเก็บข้อมูล มากขึ้น ขนาดของข้อมูลใกล้เคียงกับข้อมูลจริง ความเร็วในการ แสดงภาพมีความเร็วมากขึ้น และสามารถค้นหาข้อมูลไบนารี ด้วยข้อความ แต่วิธีการนี้จะมีข้อเสียบางประการนั่นคือขนาด ข้อมูลจะมีขนาดใหญ่ขึ้นกว่าขนาดข้อมูลจริง เหมาะกับการใช้ งานในด้านข้อมูลที่เก็บเป็นไฟล์ไม่ว่าจะเป็นข้อมูลชนิด .png .jpg .doc .xls .xlsx .pdf และอื่นๆ เป็นต้น

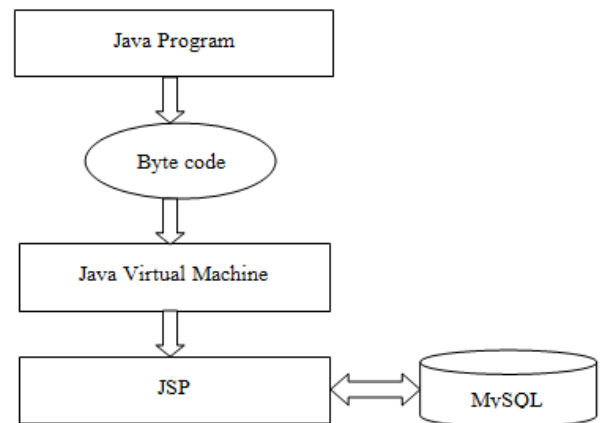
2.2 การจัดการข้อมูลบนเว็บ

งานวิจัยในนี้ มีการพัฒนาโปรแกรมด้วยภาษาจาวาและทำ การจัดเก็บข้อมูลลงฐานข้อมูล MySQL โดยเก็บข้อมูลชนิด varchar(1024) มุ่งเน้นที่จะสร้างไฟล์นามสกุล .jar เพื่อเป็น ไลบรารี (Library) กลางให้เรียกใช้ได้กับโปรเจ็คต่าง ๆ ได้มี ข้อดีคือ โปรเจ็คต่างๆ สามารถเรียกใช้งานได้ถ้ามีไฟล์ Base128.jar มีข้อเสียอยู่เล็กน้อยคือใช้งานได้เฉพาะโปรแกรมที่ สร้างด้วยภาษาจาวา

3. ระเบียบวิธีวิจัย

3.1 การออกแบบ

งานวิจัยชิ้นนี้ได้มีการออกแบบให้ทำงานอยู่บนระบบเว็บ แอปพลิเคชันโดยพัฒนาด้วยภาษา JSP (Java Server Pages) ให้ สามารถใช้งานได้กับระบบปฏิบัติการทุกแพลตฟอร์ม (Platform) โดยสร้างจาวาคลาสเพื่อเป็นไลบรารีกลางในการ เรียกใช้ ดังภาพที่ 1



ภาพที่ 1: ภาพรวมของโปรแกรม

การทำงานบนเว็บเบราว์เซอร์ โดยการเลือกภาพหรือไฟล์โดย กดปุ่มอัป โหลด ระบบจะแสดงสถิติการเปรียบเทียบการ เข้ารหัสระหว่าง Base32 Base64 และ Base128

3.2 ประสิทธิภาพของระบบ

ประสิทธิภาพของระบบทดสอบโดยสมการคำนวณหา ระยะเวลาที่ใช้ในการบันทึกข้อมูลมีดังนี้

$$t_w = t_c + t_i \quad (1)$$

เมื่อ t_w คือ เวลารับที่ก

t_c คือ เวลาเข้ารหัสและแปลงข้อมูล

t_i คือ เวลาจัดเก็บลงฐานข้อมูล

และสมการคำนวณหาเวลาที่ใช้ในการดึงข้อมูลกลับมาประกอบด้วยเวลาของการถอดรหัสรวมกับเวลาของการเรียกข้อมูลจากฐานข้อมูลดังสมการ

$$t_r = t_q + t_d \quad (2)$$

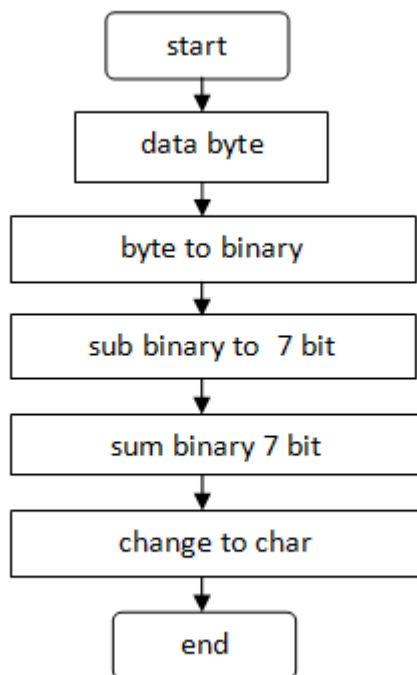
เมื่อ t_r คือ เวลาที่ใช้ในการดึงข้อมูลกลับมา

t_q คือ เวลาเรียกข้อมูลจากฐานข้อมูล

t_d คือ เวลาแปลงข้อมูลกลับ

3.3 เทคนิคและขั้นตอนวิธี

เทคนิคที่ใช้คือการแปลงข้อมูลชนิด Byte [4] เป็น binary ทำการตัด binary ครั้งละ 7 bit [6] แล้วแปลงค่าจาก binary เป็นเลขฐานสิบ และนำเลขฐานสิบไปแทนค่าด้วยอักขระของข้อมูลสตริง ดังภาพที่ 2



ภาพที่ 2: ขั้นตอนวิธีการทำงาน

3.3 การเข้ารหัส

Base128 มีการเข้ารหัสด้วยอินเด็กซ์ของข้อมูลสตริงดังตารางที่ 1

ตารางที่ 1: อินเด็กซ์ของข้อมูลสตริง

Value	Char	Value	Char	Value	Char	Value	Char	Value	Char
0	A	27	b	54	2	81]	108	ณ
1	B	28	c	55	3	82	^	109	ด
2	C	29	d	56	4	83	_	110	ค
3	D	30	e	57	5	84	`	111	ช
4	E	31	f	58	6	85	{	112	ท
5	F	32	g	59	7	86		113	ธ
6	G	33	h	60	8	87	}	114	น
7	H	34	i	61	9	88	~	115	บ
8	I	35	j	62	!	89	\	116	ป
9	J	36	k	63	#	90	"	117	ม
10	K	37	l	64	\$	91	'	118	ฟ
11	L	38	m	65	%	92	"	119	พ
12	M	39	n	66	&	93	"	120	ฟ
13	N	40	o	67	(94	ค	121	ภ
14	O	41	p	68)	95	ค	122	ม
15	P	42	q	69	*	96	"	123	ธ
16	Q	43	r	70	+	97	"	124	ร
17	R	44	s	71	,	98	จ	125	ธ
18	S	45	t	72	.	99	ฉ	126	ว
19	T	46	u	73	/	100	ช	127	ศ
20	U	47	v	74	:	101	ช		
21	V	48	w	75	;	102	ณ		
22	W	49	x	76	<	103	ญ		
23	X	50	y	77	>	104	ญ		
24	Y	51	z	78	?	105	ฐ		
25	Z	52	0	79	@	106	ชา		
26	a	53	1	80	[107	ณ		

3.4 การแปลงข้อมูล

การแปลงข้อมูลในงานวิจัยนี้ เป็นส่วนสำคัญที่สุด ที่จะช่วยลดระยะเวลาในการจัดเก็บ และช่วยลดขนาดของข้อมูลไบนารีที่ต้องจัดเก็บในฐานข้อมูล ดังตารางที่ 2

ตารางที่ 2: การแปลงข้อมูล

Bit pattern	0 1 1 0 0 0 1 0 0 1 0 1 0 1 1 1 1 0 1 1 1 0 0 1 1 0 0 1 0 0 1 0 0 0 0 1
Index	49 21 96 50 1
Base128- encoded	s v g y R

3.5 ไฟล์ทดสอบ

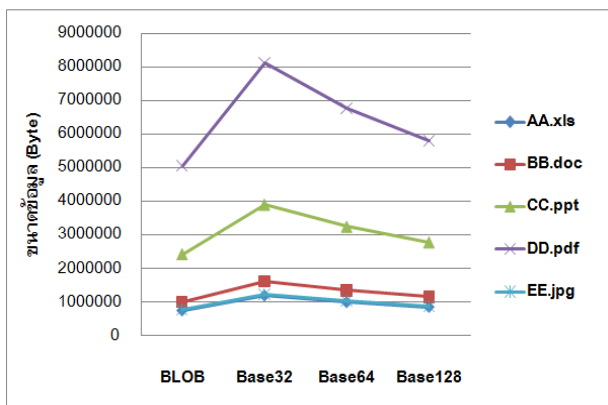
การทดสอบประสิทธิภาพของระบบใช้ไฟล์เพื่อทดสอบระบบดังแสดงในตารางที่ 3

ตารางที่ 3: ไฟล์ทดสอบ

ประเภทข้อมูล	ขนาดข้อมูล
AA.xls	726 KB
BB.doc	992 KB
CC.ppt	2374 KB
DD.pdf	4952 KB
EE.jpg	760 KB

4. ผลการทดสอบ

จากการทดสอบและใช้งานจริงกับชุดไฟล์ทดสอบดังแสดงในตารางที่ 3 ผ่านระบบที่จำลองขึ้นมาเพื่อทดสอบ โดยทำการทดสอบผ่านเว็บแอปพลิเคชัน โดยเลือกไฟล์ข้อมูล ซึ่งผลการทดสอบขั้นต้นคือการวัดขนาดของข้อมูลที่ต้องทำการจัดเก็บลงในฐานข้อมูลเปรียบเทียบกับการเข้ารหัส และแปลงเป็นข้อมูลแบบสตริงวิธีต่างๆ โดยได้ผลดังแสดงในภาพที่ 3 โดยจะพบว่าขนาดข้อมูลที่ใช้การเข้ารหัสและแปลงแบบ Base128 มีขนาดมากกว่าการจัดเก็บข้อมูลแบบ BLOB เพียงเล็กน้อย



ภาพที่ 3: ขนาดของข้อมูลที่ใช้ในการจัดเก็บ

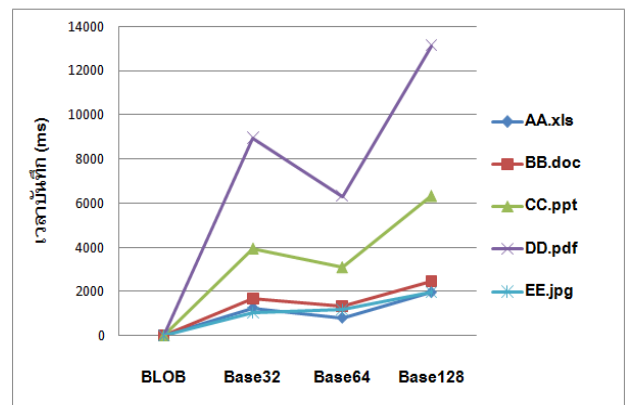
ระบบจะทำการเก็บข้อมูลเวลาในการเข้ารหัส เวลาในการจัดเก็บข้อมูลลงฐานข้อมูล เวลาของการดึงข้อมูลกลับและเวลาในการแสดงไฟล์ข้อมูล มีตัวแปรในการทดสอบคือ การเลือก

ภาพที่จะทดสอบ แสดงผลการทดสอบโดยแสดงประสิทธิภาพในการลดขนาดของข้อมูล ความเร็วในการเข้ารหัสข้อมูล ความเร็วในการจัดเก็บข้อมูล ความเร็วในการแสดงภาพหรือไฟล์ ความเร็วในการดึงข้อมูลกลับดังรายละเอียดต่อไปนี้

4.1 ส่วนของการบันทึก

ผลการทดสอบที่ได้ในด้านของระยะเวลาที่ใช้ในการบันทึกข้อมูลดังสมการที่ (1)

ผลที่ได้จากการทดสอบแสดงได้ดังภาพที่ 4 เมื่อนำผลมาเปรียบเทียบกันแล้ว จะเห็นได้ว่า วิธีการของ Base32 Base 64 ที่แกน X ใช้เวลาในการจัดเก็บน้อยกว่า Base128 เนื่องจากขั้นตอนในการเข้ารหัสของ Base128 มีขั้นตอนที่ซับซ้อนกว่า Base32 และ Base64 และการบันทึกข้อมูลแบบ BLOB ใช้เวลาน้อยสุดเนื่องจาก $t_c = 0$



ภาพที่ 4: เวลาที่ใช้ในการแปลงและบันทึกข้อมูลลงฐานข้อมูล

จากผลการทดสอบในด้านเวลาของการเก็บบันทึกข้อมูลเมื่อเปรียบเทียบในแต่ละวิธีการนั้น เราสามารถเห็นความแตกต่างของแต่ละวิธีการได้อย่างชัดเจน

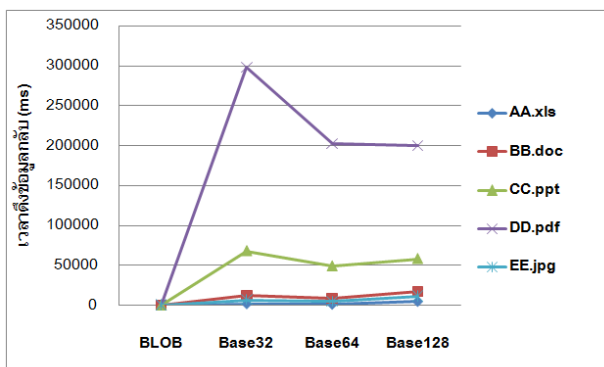
4.2 ส่วนของการแปลงข้อมูล

ส่วนของการแปลงข้อมูลไบนารีมาเป็นข้อมูลสตริงนั้น ถือเป็นส่วนสำคัญในการพัฒนางานวิจัยนี้ ซึ่งปัญหาและความยากของการพัฒนาระบบการเก็บบันทึกข้อมูลไบนารีชนิดต่าง ๆ ให้มีประสิทธิภาพมากยิ่งขึ้นนั้น โดยมุ่งเน้นในเรื่องการลดขนาดของความจุข้อมูลลงฐานข้อมูลเป็นลำดับแรก

จากผลการทดสอบการแปลงข้อมูลไบนารีของงานวิจัยนี้ เมื่อเทียบกับวิธีการ Base32 โดยเฉลี่ยจะมีขนาดลดลง 30% ถ้าเทียบกับวิธีการ Base64 โดยเฉลี่ยจะมีขนาดลดลง 20% และถ้าเทียบกับการเก็บข้อมูลแบบ BLOB เฉลี่ยแล้วจะมีขนาดเพิ่มขึ้นจากเดิม 14% ถือว่ามีขนาดใกล้เคียงกันกับขนาดจริงของไฟล์ จะเห็นได้ว่า การพัฒนาให้ระบบสามารถย่อขนาดของข้อมูลให้ลดลง หรือคงขนาดของข้อมูลเดิมไว้นั้น เป็นส่วนที่กระทำได้ยากมาก ซึ่งทำให้พื้นที่ในการเก็บบันทึกข้อมูลไบนารีต่าง ๆ ลงในฐานข้อมูลมีขนาดใกล้เคียงกับขนาดจริงให้มากที่สุด แต่อย่างไรก็ตามด้วยวิธีการยังมีส่วนที่ต้องปรับปรุงต่อไป

4.3 ส่วนของการดึงข้อมูลกลับมา

ในส่วนของการดึงข้อมูลกลับมาประกอบด้วยเวลาของการถอดรหัสรวมกับเวลาของการเรียกข้อมูลจากฐานข้อมูล ดังสมการที่ (2) ผลที่ได้จากการทดสอบแสดงได้ดังภาพที่ 5



ภาพที่ 5: เวลาที่ใช้ในการดึงและแปลงข้อมูลกลับมา

จากภาพแสดงให้เห็นว่าเวลาในการดึงข้อมูลของวิธีการ BLOB ใช้เวลาน้อยสุดเนื่องจาก $t_d = 0$ ส่วนวิธีการ Base128 ใช้เวลาในการดึงข้อมูลกลับมาได้ใกล้เคียงกับเวลาของ Base64 และบางชนิดไฟล์ภาพใช้เวลาน้อยกว่า เช่น ชนิด .pdf เป็นต้น

5. สรุปผลและเสนอแนะ

งานวิจัยนี้ได้ทำการแก้ปัญหาการลดขนาดของข้อมูลไบนารีหรือไฟล์ข้อมูลเมื่อจัดเก็บข้อมูลในภาพแบบสตริง โดยใช้วิธีสร้างคลาส Base128 ซึ่งได้ผลลัพธ์จากการลดขนาดของข้อมูลอยู่ที่ 30% เมื่อเทียบกับ Base64 ดีขึ้นจากเดิมเมื่อคิดในด้านการบันทึกข้อมูล ลดความเสี่ยงข้อมูลสูญหายระหว่างการบันทึก

ข้อมูล เนื่องจากเก็บข้อมูลครั้งเดียวในภาพแบบสตริงแตกต่างจากวิธีเดิมที่อ่านค่าการเก็บข้อมูลที่ละไบต์ ซึ่งระหว่างการอ่านค่าการเก็บที่ละไบต์อาจเกิดการสูญหาย (loss) ของข้อมูลระหว่างการจัดเก็บได้ทำให้ข้อมูลที่เก็บลงฐานข้อมูลเกิดการเสียหายได้

งานวิจัยที่จะพัฒนาต่อไปคือการทดสอบประสิทธิภาพการค้นหาข้อมูลในระบบฐานข้อมูลที่มีการแปลงเป็นข้อมูลสตริงด้วยรหัส BASE64 BASE128 ซึ่งรวมทั้งวิธีการเข้ารหัสข้อมูลหลังทำการแปลงและเข้ารหัสแล้ว [5],[7] ก่อนการจัดเก็บข้อมูล

เอกสารอ้างอิง

- [1] Mackenzie, Charles E, "ASCII," [Online] 1980. Available from : <http://en.wikipedia.org/wiki/ASCII> [2013, December 10].
- [2] S. Josefsson, "The Base16, Base32, and Base64 Data Encodings," [Online] 2006. Available from : <http://www.ietf.org/rfc/rfc4648.txt>[2013,December 15].
- [3] Dev , "base64 Image ," [Online] 2006. Available from : <http://base64image.org/> [2013,January 15].
- [4] Changeset 5814, "Base64 utility methods ," section 6.2, [Online] 2014. Available from : <http://www.w3.org/html/wg/drafts/html/master/webappapis.html#atob> [2013,December 15].
- [5] Congfu Xu, Yafang Chen, Kevin Chiew, " An Approach to Image Spam Filtering Based on Base64 Encoding and N-Gram Feature Extraction, " International Conference on Tools with Artificial Intelligence. pp. 171-176, 2010.
- [6] Zhenxing Liu , Lu Liu , Richard Hill1, Yongzhao Zhan, "Base62x: An Alternative Approach to Base64 for non-Alphanumeric Characters," Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD).pp. 2667-2670, 2011.
- [7] Gurpreet Singh, Supriya, " Modified Vigenere Encryption Algorithm and its Hybrid Implementation with Base64 and AES ," International Conference on Advanced Computing, Networking and Security,pp. 232-237, 2013.

การลดมิติข้อมูลด้วยวิธีการคัดเลือกคุณลักษณะข้อมูลแบบหลายลำดับชั้น ด้วยเทคนิค Triple Linear SVM Weights (TLiSVM) Multi-layers Feature Selection Method Using TLiSVM for Dimensionality Reduction

วิภาวรรณ บัวทอง (Wipawan Buathong)¹ และพยุ่ง มีสังข์ (Phayung Meesad)²

¹สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏภูเก็ต

¹ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

²ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹wipawan.buathong@gmail.com, ²pym@kmutnb.ac.th

บทคัดย่อ

บทความวิจัยนี้ได้นำเสนอวิธีการลดมิติข้อมูลด้วยวิธีการคัดเลือกคุณลักษณะข้อมูลแบบหลายลำดับชั้นด้วยเทคนิค Triple Linear SVM Weight (TLiSVM) เพื่อใช้กับตัวจำแนกประเภทข้อมูลแบบซัพพอร์ตเวกเตอร์แมชชีน สำหรับการทำให้เหมือนข้อมูลแบบจำแนกประเภทกับข้อมูลแบบสองคลาสโดยได้ทำการศึกษาเปรียบเทียบประสิทธิภาพกับวิธีการลดมิติข้อมูลด้วยวิธีการคัดเลือกคุณลักษณะข้อมูลแบบลำดับชั้นเดียวด้วยเทคนิค Linear SVM Weight (LiSVM) และวิธีการลดมิติข้อมูลด้วยวิธีการคัดเลือกคุณลักษณะข้อมูลแบบหลายลำดับชั้นด้วยเทคนิค Double Linear SVM Weight (DLiSVM) โดยทดสอบกับชุดข้อมูล DLBCL, Duke Breast-Cancer และ Leukemia ผลการศึกษาพบว่า วิธีการลดมิติข้อมูลแบบหลายลำดับชั้นด้วยเทคนิค TLiSVM มีประสิทธิภาพในการลดมิติข้อมูลได้ดีกว่าเทคนิค LiSVM และ DLiSVM กล่าวคือ ให้ค่าความถูกต้องของแบบจำลองการเรียนรู้ของเครื่องในการจำแนกประเภทข้อมูลอยู่ที่ 100% สำหรับทุกชุดข้อมูลและสามารถลดจำนวนคุณลักษณะข้อมูลจากทั้งหมด 7,070 คุณลักษณะ เหลือเพียง 16 คุณลักษณะข้อมูลสำหรับชุดข้อมูล DLBCL ส่วนชุดข้อมูล Duke Breast-Cancer และ Leukemia สามารถลดจำนวนคุณลักษณะข้อมูลเหลือเพียง 11 คุณลักษณะข้อมูลจากจำนวน 7,129 คุณลักษณะข้อมูล

คำสำคัญ: การลดมิติข้อมูล LiSVM DLiSVM TLiSVM การจัดลำดับมิติข้อมูลแบบหลายลำดับชั้น

Abstract

This article proposed Triple Linear SVM Weight

(TLiSVM) as an alternative technique for dimensionality reduction with a Support Vector Machine (SVM) classifier on a two-class dataset. The efficiency of TLiSVM was compared with two chosen techniques, including Linear SVM Weight (LiSVM) and Double Linear SVM Weight (DLiSVM). Three datasets, including DLBCL, Duke Breast-Cancer and Leukemia, were used for the experiment. The proposed technique was discovered to be more efficient than using either LiSVM or DLiSVM for dimensionality reduction. The accuracy rate could reach 100 percent in all experimental datasets with the same consistency of dimensionality reduction. While the dimensional data of DLBCL could be downsized from 7,070 to 16 attributes, the dimensional data of Duke Breast-Cancer and Leukemia could be downsized from 7,129 to 11 attributes.

Keywords: Dimensionality Reduction, LiSVM, DLiSVM, TLiSVM, Multi-Layers Feature Selection.

1. บทนำ

ปัจจุบันการทำเหมืองข้อมูลนิยมทำกับข้อมูลขนาดใหญ่ที่มีปริมาณข้อมูลมากเพียงพอเพื่อให้ผลลัพธ์ที่ได้มีความน่าเชื่อถือจึงมักประสบกับปัญหามิติของข้อมูลที่มากเกินไปเป็นประจำซึ่งส่งผลให้การทำให้เหมือนข้อมูลขาดประสิทธิภาพ เช่น ใช้ระยะเวลานานในการสร้างแบบจำลองการเรียนรู้ของเครื่อง ใช้ทรัพยากรในการประมวลผลมาก ความถูกต้องของแบบจำลองการเรียนรู้ของเครื่องที่ได้อยู่ในระดับต่ำและแบบจำลองการ

เรียนรู้ของเครื่องมีความซับซ้อนยากต่อการทำความเข้าใจ ดังนั้นจึงจำเป็นต้องมีการลดมิติข้อมูลลงการลดมิติข้อมูลเป็นกระบวนการหนึ่งในขั้นตอนการเตรียมข้อมูล [1] กล่าวคือ เป็นการทำให้ข้อมูลตั้งต้นมีขนาดลดลงโดยสูญเสียลักษณะสำคัญของข้อมูลและความถูกต้องของผลลัพธ์น้อยที่สุด เนื่องจากข้อมูลแต่ละมิติจะมีความสำคัญต่อความถูกต้องของแบบจำลองการเรียนรู้ของเครื่องในการจำแนกประเภทข้อมูลไม่เท่ากัน ด้วยเทคนิคการเลือกมิติข้อมูลที่ดียิ่งจะส่งผลทำให้สามารถเลือกมิติข้อมูลที่มีความสำคัญและสามารถใช้เป็นตัวแทนของข้อมูลทั้งหมดได้ มิติของข้อมูลที่ลดลงจะทำให้การทำเหมืองข้อมูลมีประสิทธิภาพมากขึ้น กล่าวคือ จะใช้ระยะเวลาอันน้อยลงในการสร้างแบบจำลองการเรียนรู้ของเครื่อง ใช้ทรัพยากรในการประมวลผลน้อยลง ความถูกต้องของแบบจำลองการเรียนรู้ของเครื่องเพิ่มสูงขึ้น และแบบจำลองการเรียนรู้ของเครื่องที่ได้ไม่ซับซ้อนทำให้ง่ายต่อการทำความเข้าใจ

จากปัญหาที่เกี่ยวกับการลดมิติข้อมูลโดยใช้วิธีการคัดเลือกคุณลักษณะของข้อมูล จะพบว่างานวิจัยส่วนใหญ่ต้องการลดมิติให้ได้มากและยังคงความแม่นยำไว้เหมือนเดิมหรือมากกว่าเดิม แต่การที่จะได้สิ่งเหล่านั้นจะต้องคำนึงถึงเทคนิคในการลดมิติข้อมูล รวมไปถึงรูปแบบของการคัดเลือกคุณลักษณะของข้อมูลในแบบต่างๆ และความสามารถของตัวจำแนกข้อมูลอีกด้วย ดังเช่นงานวิจัยของ [2] ที่จำแนกข้อมูลประเภทข้อมูลและสัญลักษณ์ที่มีคลาสเพียงสองคลาสได้ดีโดยใช้อัลกอริทึมของ Decision Tree แต่เมื่อจำนวนแอททริบิวต์เพิ่มขึ้นอัลกอริทึมนี้จะมีประสิทธิภาพลดลง [3-5] นำเสนอเทคนิคการลดมิติข้อมูลโดยใช้วิธีการคัดเลือกคุณลักษณะของข้อมูลกับจำนวนข้อมูลที่มีแอททริบิวต์จำนวนมากและพบว่าซัพพอร์ตเวกเตอร์แมชชีนเป็นตัวจำแนกข้อมูลที่มีประสิทธิภาพมากกว่าตัวจำแนกข้อมูลชนิดอื่นๆ ซึ่ง [6-9] รวมถึงนักวิจัยอีกหลายคนได้มีการกำหนดค่าพารามิเตอร์ที่เหมาะสมให้กับวิธีซัพพอร์ตเวกเตอร์แมชชีนทำให้สามารถเพิ่มประสิทธิภาพในการจำแนกข้อมูลได้ดีขึ้น

ทั้งนี้วิธีการลดมิติข้อมูลที่ใช้ในงานวิจัยจำนวนมากอยู่บนพื้นฐานของการพิจารณาคัดเลือกคุณลักษณะข้อมูลที่สำคัญเพียงครั้งเดียวด้วยการพิจารณาคัดเลือกเพียงลำดับชั้นเดียวและใช้เทคนิคเดียวเท่านั้นหรืออาจกล่าวได้ว่าใช้เกณฑ์เพียงเกณฑ์เดียวในการพิจารณาคัดเลือกมิติข้อมูล ซึ่งอาจทำให้เกิดปัญหา

มิติข้อมูลที่มีความสำคัญไม่ได้รับการคัดเลือก เพื่อแก้ไขปัญหา มิติข้อมูลที่มีความสำคัญไม่ได้รับการคัดเลือกและปัญหาการจัดเรียงลำดับคุณลักษณะจากการใช้เกณฑ์เพียงเกณฑ์เดียวในการพิจารณาคัดเลือกคุณลักษณะ ดังนั้นผู้วิจัยจึงมีแนวคิดที่จะเพิ่มเกณฑ์การพิจารณาในการคัดเลือกคุณลักษณะข้อมูลที่มีความสำคัญโดยใช้วิธีการลดมิติข้อมูลแบบเพิ่มลำดับชั้นในการพิจารณาคัดเลือกมิติข้อมูล เพื่อเพิ่มประสิทธิภาพของวิธีการลดมิติและการทำเหมืองข้อมูล

2. ทฤษฎีที่เกี่ยวข้อง

2.1 การลดมิติข้อมูลประเภทการคัดเลือกคุณลักษณะ

วิธีการลดมิติข้อมูลสามารถจำแนกออกได้เป็น 2 ประเภท คือ การคัดเลือกคุณลักษณะ (Feature Selection) และการสกัดคุณลักษณะ (Feature Extraction) ในงานวิจัยนี้จะกล่าวถึงเฉพาะการคัดเลือกคุณลักษณะเท่านั้น

การคัดเลือกคุณลักษณะเป็นกระบวนการพิจารณาคัดเลือกคุณลักษณะของข้อมูลจำนวนหนึ่งจากคุณลักษณะของข้อมูลทั้งหมดที่สามารถใช้เป็นตัวแทนของข้อมูลทั้งหมดได้ หากข้อมูลมีคุณลักษณะที่มากเกินไปจนเกินไป เมื่อนำข้อมูลไปใช้ในการทำเหมืองข้อมูลจะทำให้ผลลัพธ์ที่ได้มีประสิทธิภาพด้อยลง [10-11] ขั้นตอนการทำงานของวิธีการคัดเลือกคุณลักษณะ เริ่มจากการนำชุดข้อมูลที่มีมิติสูงมาสร้างเป็นเซตของคุณลักษณะข้อมูลตั้งต้น ซึ่งถ้าคุณลักษณะของข้อมูลตั้งต้นมีจำนวน n คุณลักษณะ จะทำให้จำนวนสมาชิกของเซตที่ได้จะมีขนาด 2^n ตัว จากนั้นจะทำการค้นหาคุณลักษณะข้อมูลที่มีโอกาสได้รับการคัดเลือกจากเซตของคุณลักษณะข้อมูลตั้งต้นเพื่อสร้างเซตย่อยซึ่งวิธีการค้นหาคุณลักษณะข้อมูลที่ต้องการมีหลายวิธี เช่น Exhaustive Search, Heuristic Search หรือ Random Search

2.2 เทคนิคการคัดเลือกคุณลักษณะแบบ LiSVM

LiSVM เป็นวิธีการคัดเลือกคุณลักษณะข้อมูลแบบลำดับชั้นเดียวที่อาศัยหลักการของซัพพอร์ตเวกเตอร์แมชชีนมาช่วยในการคำนวณค่าน้ำหนักของแต่ละคุณลักษณะข้อมูล เพื่อใช้พิจารณาคัดเลือกคุณลักษณะข้อมูลที่ให้ค่าน้ำหนักมากตามที่ต้องการ จากการศึกษาวิจัยพบว่า LiSVM สามารถประยุกต์ใช้ได้กับตัวจำแนกประเภทข้อมูล (Classifier) [12] ทุกตัว โดยมีอัลกอริทึมดังภาพที่ 1

จากอัลกอริธึมดังภาพที่ 1 เมื่อใส่อินพุตที่เป็นชุดข้อมูลฝึกฝนเข้าไปประมวลผลแล้วจะได้เอาต์พุตเป็นคุณลักษณะข้อมูลที่จัดเรียงลำดับตามค่าน้ำหนักที่คำนวณได้ ซึ่งประกอบด้วยขั้นตอน 3 ขั้นตอนดังต่อไปนี้

Algorithm 1 Feature Ranking Based on Linear SVM Weights

Input : Training sets, $(X_i, Y_i), i=1, \dots, m.$

Output : Sorted feature ranking list.

1. Use grid search to find the best parameter C.
2. Train aL2-loss linear SVM model using the best C.
3. Sort the features according to the absolute values of weights in the model.

ภาพที่ 1: อัลกอริธึมของ Linear SVM Weight

ขั้นตอนแรกจะใช้การค้นหาแบบกริด (grid) เพื่อค้นหาค่าพารามิเตอร์ C ที่ดีที่สุดสำหรับชุดข้อมูล จากนั้นในขั้นตอนที่สองจะทำการสร้างโมเดลด้วยซัพพอร์ตเวกเตอร์แมชชีนที่ใช้เคอร์เนลฟังก์ชันแบบเชิงเส้นตามสมการที่ 1 เพื่อหาค่าระยะขอบที่ห่างที่สุดและใช้ฟังก์ชัน L2-loss [13-14] ดังสมการที่ 2 เพื่อปรับค่า Slack Vector กับชุดข้อมูลโดยใช้พารามิเตอร์ C ที่ดีที่สุดที่ได้จากขั้นตอนที่ 1 เพื่อคำนวณหาค่าน้ำหนักของแต่ละคุณลักษณะ ในขั้นตอนสุดท้ายจะทำการเรียงลำดับคุณลักษณะข้อมูลโดยพิจารณาจากค่าสัมบูรณ์ของค่าน้ำหนักที่ได้

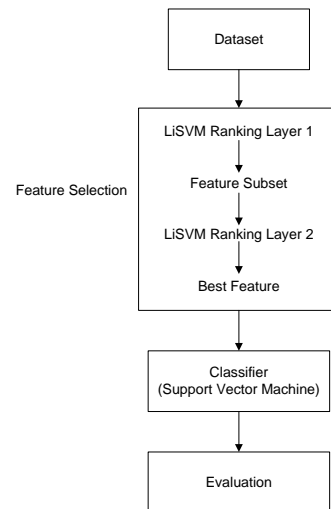
$$\min_{w,b} \frac{1}{2}w^T w + C \sum_{i=1}^l \xi(w, b; x_i, y_i) \quad (1)$$

$$\xi(w, b; x_i, y_i) = \max(1 - y_i(w^T x_i + b), 0^2) \quad (2)$$

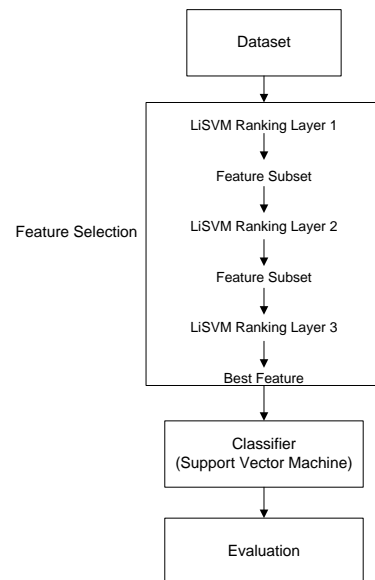
2.3 เทคนิคการคัดเลือกคุณลักษณะแบบ DLiSVM

DLiSVM เป็นวิธีการคัดเลือกคุณลักษณะข้อมูลแบบหลายลำดับชั้นที่อาศัยหลักการของซัพพอร์ตเวกเตอร์แมชชีนมาช่วยในการคำนวณหาค่าน้ำหนักของแต่ละคุณลักษณะข้อมูล เช่นเดียวกับ LiSVM โดยการทำงานของ DLiSVM จะเริ่มจากการพิจารณาคัดเลือกคุณลักษณะข้อมูลจากชุดข้อมูลตั้งต้นด้วยเทคนิคแบบ LiSVM เพื่อให้ได้คุณลักษณะข้อมูลจำนวนหนึ่งที่มีมิติข้อมูลลดลงซึ่งถือเป็นรอบแรกในการพิจารณาคัดเลือก จากนั้นจะนำคุณลักษณะข้อมูลที่ได้ข้างต้นมาพิจารณาคัดเลือกคุณลักษณะข้อมูลใหม่ในรอบที่ 2 ด้วยเทคนิคแบบ LiSVM อีกครั้งหนึ่งเพื่อให้ได้คุณลักษณะข้อมูลตามที่ต้องการ เมื่อได้คุณลักษณะข้อมูลตามจำนวนที่ต้องการแล้วก็จะสามารถนำเข้าสู่กระบวนการทำเหมืองข้อมูลต่อไปดังภาพที่ 2 จากการ

ศึกษาวิจัยพบว่า DLiSVM มีประสิทธิภาพในการลดมิติข้อมูลดีกว่า LiSVM [15]



ภาพที่ 2: การคัดเลือกคุณลักษณะด้วยวิธี DLiSVM



ภาพที่ 3: การคัดเลือกคุณลักษณะด้วยวิธี TLiSVM

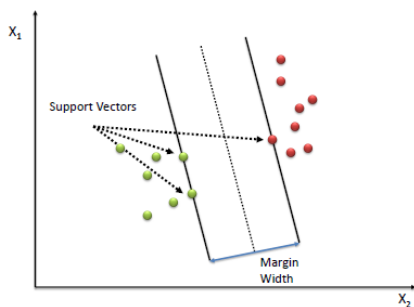
2.4 เทคนิคการคัดเลือกคุณลักษณะแบบ TLiSVM

TLiSVM เป็นวิธีการคัดเลือกคุณลักษณะข้อมูลแบบหลายลำดับชั้นเช่นเดียวกับ DLiSVM ที่อาศัยหลักการของซัพพอร์ตเวกเตอร์แมชชีนมาช่วยในการคำนวณหาค่าน้ำหนักของแต่ละคุณลักษณะข้อมูล โดยการทำงานของ TLiSVM จะเริ่มจากการพิจารณาคัดเลือกคุณลักษณะข้อมูลจากชุดข้อมูลตั้งต้นด้วยเทคนิคแบบ LiSVM เพื่อให้ได้คุณลักษณะข้อมูลจำนวนหนึ่งที่มีมิติข้อมูลลดลงซึ่งถือเป็นรอบแรกในการพิจารณาคัดเลือก จากนั้นจะนำคุณลักษณะข้อมูลที่ได้ข้างต้นมาพิจารณาคัดเลือกคุณลักษณะข้อมูลใหม่ในรอบที่ 2 ด้วยเทคนิคแบบ

LiSVM อีกครั้งหนึ่งเพื่อให้ได้คุณลักษณะข้อมูลจำนวนหนึ่งที่มีมิติข้อมูลลดลง จากนั้นจะนำคุณลักษณะข้อมูลที่ได้ในรอบที่ 2 มาพิจารณาคัดเลือกคุณลักษณะข้อมูลใหม่ในรอบที่ 3 ด้วยเทคนิคแบบ LiSVM อีกครั้งหนึ่ง เพื่อให้ได้คุณลักษณะข้อมูลตามที่ต้องการ เมื่อได้คุณลักษณะข้อมูลตามจำนวนที่ต้องการแล้วก็จะสามารถนำเข้าสู่กระบวนการทำเหมืองข้อมูลต่อไปดังภาพที่ 3

2.5 ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine)

ซัพพอร์ตเวกเตอร์แมชชีน เป็นเทคนิคที่ใช้ในการสร้างแบบจำลองการเรียนรู้ของเครื่องเพื่อจำแนกประเภทข้อมูลที่ได้รับค่านิยม [16] เนื่องจากให้ความถูกต้องในการจำแนกประเภทข้อมูลสูง ดังปรากฏในงานวิจัยหลายชิ้น [17-18] ซัพพอร์ตเวกเตอร์แมชชีนใช้หลักการหาระยะขอบที่กว้างที่สุด (Maximun Margin) ของระนาบการตัดสินใจ (Decision Hyper Plane) เพื่อใช้ระนาบดังกล่าวในการจำแนกประเภทข้อมูลออกจากกันซึ่งระนาบการตัดสินใจอาจมีหลายระนาบ ดังนั้นจำเป็นต้องเลือกระนาบการตัดสินใจที่ดีที่สุดเพียงระนาบเดียวด้วยการพิจารณาจุดที่อยู่บนระนาบและเลือกระนาบการตัดสินใจที่ทำให้เกิดระยะห่างหรือระยะขอบของข้อมูลแต่ละกลุ่มให้มากที่สุดดังภาพที่ 4 ระนาบการตัดสินใจที่ใช้จำแนกประเภทข้อมูลอาจมีลักษณะเป็นเส้นตรงหรือไม่เป็นเส้นตรงก็ได้



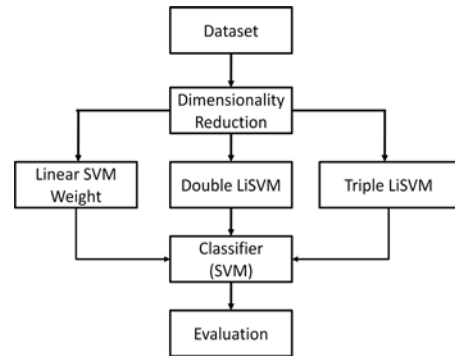
ภาพที่ 4: แสดงการเลือกระนาบการตัดสินใจที่ดีที่สุดในการจำแนกประเภทข้อมูล

โดยปกติหากข้อมูลตัวอย่างที่นำมาใช้ในการสร้างแบบจำลองการเรียนรู้ของเครื่องด้วยซัพพอร์ตเวกเตอร์แมชชีนไม่สามารถหาระนาบการตัดสินใจที่เหมาะสมได้

3. วิธีดำเนินการวิจัย

ขั้นตอนที่ใช้ในการดำเนินงานศึกษาวิจัยในครั้งนี้มีทั้งหมด 3 ขั้นตอน คือ 1) การคัดเลือกและเตรียมชุดข้อมูล 2) เครื่องมือ

และเทคนิคที่ใช้ในการลดมิติข้อมูล และ3) การวัดและประเมินผล ซึ่งรายละเอียดของแต่ละขั้นตอนมีดังภาพที่ 5



ภาพที่ 5: แสดงขั้นตอนการดำเนินงานวิจัย

3.1 การคัดเลือกและเตรียมชุดข้อมูล

การศึกษาวิจัยครั้งนี้ใช้ชุดฐานข้อมูล <http://orange.Biolab.si/datasets.php> จำนวน 3 ชุดข้อมูล โดยทุกชุดข้อมูลที่เลือกจะไม่มีค่าที่หายไป (Missing Value) ดังตารางที่ 1

ตารางที่ 1 : ชุดข้อมูลที่ใช้ในการทดลอง

Dataset	Features	instances	Number of Class
DLBCL	7,070	77	2
Duke-Breast-Cancer	7,129	44	2
Leukemia	7,129	72	2

3.2 เครื่องมือและเทคนิคที่ใช้ในการลดมิติข้อมูล

ใช้ซอฟต์แวร์ Orange Canvas version 2.72 ทำการปรับลดมิติข้อมูลด้วยวิธีการคัดเลือกคุณลักษณะที่สำคัญแบบ LiSVM, DLiSVM และ TLiSVM และใช้โปรแกรม Matlab เวอร์ชัน 2012b ในการสร้างแบบจำลองการเรียนรู้ของเครื่องเพื่อจำแนกประเภทข้อมูลโดยใช้เทคนิคซัพพอร์ตเวกเตอร์แมชชีน ใช้ฟังก์ชัน Linear เคอร์เนลและนำข้อมูลที่ปรับลดมิติแล้วมาสร้างแบบจำลองการเรียนรู้ของเครื่อง (Machine Learning Model)

3.3 วัดและประเมินผล

การวัดและประเมินรวมถึงการเปรียบเทียบประสิทธิภาพของโมเดลในการลดมิติข้อมูลแต่ละแบบในงานวิจัยนี้ใช้วิธีการ 10-Fold Cross Validation เป็นวิธีการทดสอบประสิทธิภาพของแบบจำลองการเรียนรู้ของเครื่องด้วยการแบ่งชุดข้อมูลออกเป็น k ชุดเท่าๆ กัน โดยใช้ข้อมูลจำนวน k-1 ชุด เพื่อทำการสร้างแบบจำลองการเรียนรู้ของเครื่องส่วนข้อมูลอีก 1 ชุดที่เหลือจะใช้สำหรับทดสอบความถูกต้องและจะทำซ้ำจนกว่าข้อมูลทุกชุด

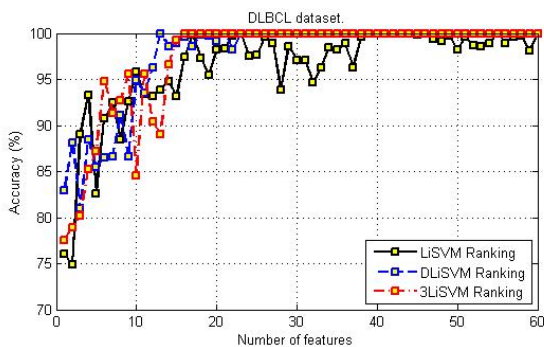
จะถูกใช้ทดสอบแบบจำลองการเรียนรู้ของเครื่อง แล้วนำค่าความถูกต้องหรือความผิดพลาดของแต่ละรอบมารวมกันและหาค่าเฉลี่ยเพื่อเป็นค่าสะท้อนประสิทธิภาพของการเรียนรู้ สำหรับการประเมินประสิทธิภาพของโมเดลใช้วิธีการวัดค่าความถูกต้อง ดังสมการที่ 3

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

เมื่อกำหนดให้ N คือ จำนวนคลาส, TP คือ ค่าของ True Positive และ TN คือ ค่าของ True Negative

4. ผลการดำเนินงาน

ผลจากการทดลองกับข้อมูล DLBCL โดยนำเทคนิคทั้งสามเทคนิคเปรียบเทียบกับกันได้แก่ LiSVM, DLiSVM และ TLiSVM ดังภาพที่ 6

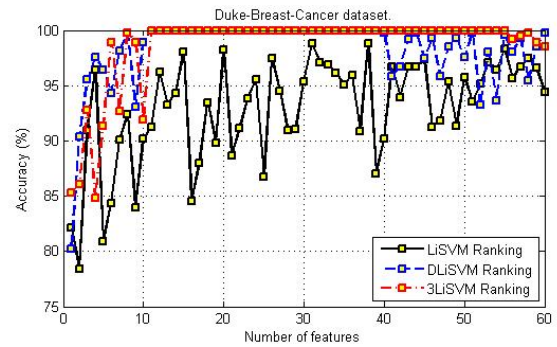


ภาพที่ 6: ผลลัพธ์ของเทคนิคต่างๆ กับข้อมูล DLBCL

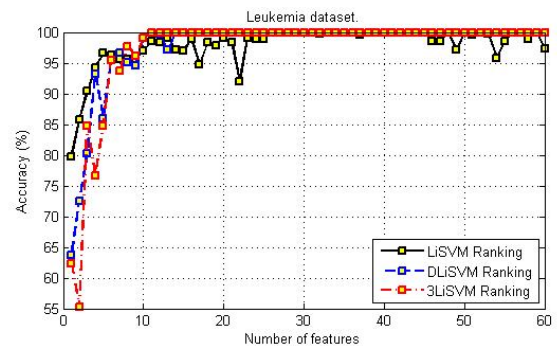
จากภาพที่ 6 พบว่าค่าความถูกต้องสูงสุดของ LiSVM อยู่ที่ 100% ซึ่งลดคุณลักษณะข้อมูลจาก 7,070 คุณลักษณะเหลือเพียง 39 คุณลักษณะข้อมูล ส่วน DLiSVM ค่าความถูกต้องสูงสุดอยู่ที่ 100% สามารถลดคุณลักษณะข้อมูลเหลือเพียง 21 คุณลักษณะข้อมูล สำหรับ TLiSVM ค่าความถูกต้องสูงสุดอยู่ที่ 100% สามารถลดคุณลักษณะข้อมูลได้ที่จำนวน 16 คุณลักษณะข้อมูล นอกจากนี้ยังพบอีกว่า DLiSVM จะมีค่าความถูกต้องคงที่ตลอดตั้งแต่จำนวนคุณลักษณะข้อมูล 23 คุณลักษณะข้อมูล ไปจนถึง 60 คุณลักษณะข้อมูล ในขณะที่เทคนิคแบบ TLiSVM 16 คุณลักษณะข้อมูล ไปจนถึง 60 คุณลักษณะข้อมูลจะให้ค่าความถูกต้องคงที่ตลอดตั้งแต่จำนวนคุณลักษณะข้อมูล

จากภาพที่ 7 สามารถอธิบายผลการทดลองได้ว่าจากการทดลองกับชุดข้อมูล Duke-Breast-Cancer จะให้ค่าความถูกต้องสูงสุดของ LiSVM อยู่ที่ 98.77% ลดคุณลักษณะข้อมูลจาก 7,129 คุณลักษณะเหลือเพียง 31 คุณลักษณะข้อมูล ส่วน DLiSVM และ TLiSVM ค่าความถูกต้องสูงสุดอยู่ที่ 100%

สามารถลดคุณลักษณะข้อมูลเหลือเพียง 11 คุณลักษณะข้อมูล นอกจากนี้ยังพบว่า TLiSVM จะมีค่าความถูกต้องคงที่ตลอดตั้งแต่จำนวนคุณลักษณะข้อมูล 11 คุณลักษณะข้อมูล ไปจนถึง 55 คุณลักษณะข้อมูล ในขณะที่เทคนิคแบบ DLiSVM จะให้ค่าความถูกต้องคงที่ตลอดตั้งแต่จำนวนคุณลักษณะข้อมูลที่ 11 คุณลักษณะข้อมูล ไปจนถึง 40 คุณลักษณะข้อมูล



ภาพที่ 7: ผลลัพธ์ของเทคนิคต่างๆ กับข้อมูล Duke-Breast-Cancer



ภาพที่ 8: ผลลัพธ์ของเทคนิคต่างๆ กับข้อมูล Leukemia

ภาพที่ 8 สามารถอธิบายการทดลองที่กระทำกับชุดข้อมูล Leukemia โดยใช้เทคนิคต่างๆ ได้ผลการทดลองดังนี้ ค่าความถูกต้องสูงสุดของ LiSVM อยู่ที่ 100% ซึ่งลดคุณลักษณะข้อมูลจาก 7,129 คุณลักษณะเหลือเพียง 26 คุณลักษณะข้อมูล ส่วน DLiSVM ค่าความถูกต้องสูงสุดอยู่ที่ 100% สามารถลดคุณลักษณะข้อมูลเหลือเพียง 14 คุณลักษณะข้อมูลและมีค่าความถูกต้องคงที่ตลอดไปจนถึง 60 คุณลักษณะข้อมูล สำหรับ TLiSVM ค่าความถูกต้องสูงสุดอยู่ที่ 100% สามารถลดคุณลักษณะข้อมูลเหลือเพียง 11 คุณลักษณะข้อมูล และมีค่าความถูกต้องคงที่ตลอดไปจนถึง 60 คุณลักษณะข้อมูล

ตารางที่ 2 : สรุปผลการทดลองชุดข้อมูลกับเทคนิคต่างๆ

เทคนิค \ ชุดข้อมูล	LiSVM		DLiSVM		TLiSVM	
	Accuracy	Features	Accuracy	Features	Accuracy	Features
DLBCL	100%	39	100%	21	100%	16
Duke-Breast-Cancer	98.77%	31	100%	11	100%	11
Leukemia	100%	26	100%	14	100%	11

5. สรุปและอภิปรายผล

ด้วยปัญหาเทคนิคการคัดเลือกคุณลักษณะข้อมูลแบบดั้งเดิมยังมีประสิทธิภาพต่ำ งานวิจัยนี้เสนอเทคนิคใหม่โดยใช้เทคนิคการคัดเลือกคุณลักษณะข้อมูลแบบหลายลำดับชั้นโดยใช้เทคนิค TLiSVM ผลการวิจัยพบว่าเทคนิคใหม่ให้ค่าความถูกต้องที่สูงและคงที่ ดีกว่าเทคนิคการคัดเลือกคุณลักษณะข้อมูลแบบลำดับชั้นเดียว รวมไปถึงจำนวนคุณลักษณะข้อมูลที่ลดลงของเทคนิคการคัดเลือกคุณลักษณะข้อมูลแบบหลายลำดับชั้นสามารถลดจำนวนคุณลักษณะได้มากกว่าเทคนิคการคัดเลือกคุณลักษณะข้อมูลแบบลำดับชั้นเดียวโดยใช้เทคนิค Linear SVM Weight และแบบหลายลำดับชั้นซึ่งใช้เทคนิค Double Linear SVM Weight

เอกสารอ้างอิง

- [1] Tan P. N., Steinbach, M., and Vipin, K. *Introduction to Data Mining*. [n.p.], United State of America : Addison Wesley, 2005.
- [2] Dharminder kumar and Suman, "Performance Analysis of Various Data Mining Algorithm: A Review." *International Journal of Computer Applications*, 32(6), Oct 2011 : 9-15.
- [3] นิเวศ จิระวิจิตรชัย, ปริญา สวงนิตย์ และพวง มีสัง, "การพัฒนาประสิทธิภาพการจัดหมวดหมู่เอกสารภาษาไทยแบบอัตโนมัติ." *การประชุมวิชาการระดับประเทศด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 6*, 28-29 ตุลาคม, 2553.
- [4] กัทราวุฒิ แสงศิริ, ศจีมาง ณ วิเชียร และพวง มีสัง, "การคัดแยกประเภทของมะเร็งเม็ดเลือดขาวโดยใช้วิธีการจัดอันดับร่วมกับเทคนิคซัพพอร์ตเวกเตอร์แมชชีน." *วารสารวิจัยมหาวิทยาลัยขอนแก่น*. ปีที่ 2 ฉบับที่ 10 เมษายน-มิถุนายน พ.ศ.2553 หน้า 10-17.
- [5] วงกต ศรีอุไร, พวง มีสัง และชุธาติ หฤไชยะศักดิ์. "การเตรียมฟีเจอร์บนพื้นฐานแบบจำลองหัวข้อสำหรับการจำแนกหมวดหมู่ของเอกสาร." *การประชุมวิชาการระดับประเทศด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 5*, 23-23 พฤษภาคม, 2552.
- [6] Yin-Wen Chang and Chih-Jen Lin. "Feature Ranking Using Linear SVM." *JMLR Workshop & Conference Proceeding*. 3(2008), 53-64.
- [7] Buathong, W., Meesad, P. 2013. "Enhancing the Efficiency of Dimensionality Reduction using a Combined Linear SVM Weight with ReliefF Feature Selection Method" In: *IC2IT 2013*, pp. 125-134.
- [8] Lei Yu, Huan Liu. "Feature Selection for High Dimensional Data: A Fast Correlation-Based Filter Solution." *Proceedings of the Twentieth International Conference on Machine Learning (ICML-2003)*, Washington DC, 2003.
- [9] วิภาวรรณ บัวทอง. 2555. "การเปรียบเทียบประสิทธิภาพของเทคนิคการลดมิติข้อมูลด้วยวิธีการจัดอันดับแบบ Information Gain, Gain Ratio และ Linear SVM Weight." The 4th National Conference on Applied Computer Technology and Information Systems (ACTIS2012). Songkhla, THAILAND. (September 21-22, 2012) : 185-189
- [10] Jing Bi, Kun Zhang, Xiaojing Cheng. "Intrusion Detection Base on RBF Neural Network." In *International Symposium on Information Engineering and Electronic Commerce*. (2009) : 362-365.
- [11] M. Dash, H. Liu. "Feature Selection for Classification" *Intelligent Data Analysis I*, (1997) : 131-156.
- [12] Guyon, I., Weston, J., Barnhill, S., & Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Machine learning*, 46(1-3), 389-422.
- [13] Chang, Y. W., & Lin, C. J. (2008). Feature ranking using linear svm. *Causation and Prediction Challenge Challenges in Machine Learning*, 47(2).
- [14] Hsu, C.-w., Chang, C.-c., & Lin, C.-j. (2010). *A practical guide to support vector classification*: Technical report, Department of Computer Science and Information Engineering, National Taiwan University, Taipei.
- [15] Buathong, W., Meesad, P. (2014). "Double Linear Support Vector Machine for Dimensionality Reduction". *Research Journal of Applied Sciences*.
- [16] Huang, C.-L., Chen, M.-C., & Wang, C.-J. (2007). Credit scoring with a data mining approach based on support vector machines. *Expert Systems with Applications*, 33(4), 847-856.
- [17] Xuehua, L., & Lan, S. (2008). *Fuzzy Theory Based Support Vector Machine Classifier*. Paper presented at the Fuzzy Systems and Knowledge Discovery, 2008. FSKD'08. Fifth International Conference on.
- [18] Sweilam, N. H., Tharwat, A. A., & Abdel Moniem, N. K. (2010). Support vector machine for diagnosis cancer disease: A comparative study. *Egyptian Informatics Journal*, 11(2), 81-92.

**การคัดเลือกปัจจัยเด่นและการเปรียบเทียบประสิทธิภาพเทคนิคการแยกประเภท
สำหรับทำนายเพศของนกกกระทาจากคุณลักษณะภายนอกของไข่นกกระทา
Feature Selection and Efficiency Comparison of Classification Techniques for
Quail Gender Forecasting from the External Factors of Quail Eggs**

จักริน สุขสวัสดิ์ชื่น (Jakkarin Suksawatchon)¹ อุรีรัฐ สุขสวัสดิ์ชื่น (Ureerat Suksawatchon)²

และพงษ์พัฒน์ สิงห์ศรี (Pongpat Sing Sri)³

คณะวิทยาการสารสนเทศ มหาวิทยาลัยบูรพา

¹jakkaman@gmail.com, ²ureerat.w@gmail.com, ³pongpat.singsri@gmail.com

บทคัดย่อ

งานวิจัยนี้นำเสนอตัวแบบที่ใช้คัดเลือกปัจจัยเด่นจากปัจจัยภายนอกของไข่นกกระทาเพื่อใช้ในการทำนายเพศของนกกกระทา โดยใช้ 2 วิธีการในการหาความสัมพันธ์ของปัจจัยคือ Best First Search และ Stepwise Multiple Regression Analysis และนำปัจจัยเด่นที่ทำการวิเคราะห์ได้ผ่านตัวแบบในการทำนายเพศซึ่งประกอบด้วย ตัวแบบต้นไม้ในการตัดสินใจ ซัพพอร์ตเวกเตอร์แมชชีน โครงข่ายประสาทเทียม ทฤษฎีของเบย์ ขั้นตอนวิธีการเพื่อนบ้านที่ใกล้ที่สุด และกฎการเชื่อมต่อกัน ข้อมูลไข่นกกระทาที่ใช้ในการทดลองจำนวน 661 ชุด ซึ่งทำการเก็บข้อมูลตั้งแต่ การจัดซื้อไข่งจนถึงไข่ฟักเป็นตัวและทำการคัดแยกเพศได้ จากผลการทดลองพบว่าวิธีการ Best First Search วิเคราะห์ปัจจัยเด่นได้ 2 ปัจจัย จาก 7 ปัจจัย คือความกว้างและลักษณะด้านแหลมของไข่ ซึ่งตัวแบบการทำนายแบบต้นไม้ตัดสินใจ (ADTree) ให้ค่าความถูกต้องในการแยกเพศสูงที่สุดเท่ากับ 69.74% ส่วนวิธีการ Stepwise Multiple Regression Analysis วิเคราะห์ปัจจัยเด่นได้ 2 ปัจจัยคือ น้ำหนักและลักษณะด้านแหลมของไข่ ซึ่งตัวแบบการทำนายแบบซัพพอร์ตเวกเตอร์แมชชีน (LibSVM) ให้ค่าความถูกต้องในการแยกเพศสูงที่สุดเท่ากับ 70.35%

คำสำคัญ: ปัจจัยภายนอก นกกกระทา การแยกประเภท

Abstract

This research proposes the feature selection model from the external factors of quail eggs for using to be quail gender prediction. To determine the relationship of the factors, we used Best First Search and Stepwise Multiple Regression Analysis. Then the relevant factors obtained from two factor analysis methods, were considered as the input of classifier models including Decision Tree, Support Vector Machine, Neural Network, Naïve Bayes, Nearest Neighbor, and Conjunctive Rule. The data used in the research were obtained from 661 quail eggs. We collected the data since purchasing eggs until the eggs hatch and sex segregation. The results show that by using Best First Search, there are 2 relevant factors from 7 factors that are width and shape of the spikes of eggs. By using these two relevant factors as the input, we have found that the decision tree (ADTree) gives the highest accuracy to 69.74%. For Stepwise Multiple Regression Analysis, there are two relevant factors that are weight and shape of the spikes of eggs and the support vector machine (LibSVM) gives the highest accuracy to 70.35%.

Keywords: External Factors, Quail, Classification

1. บทนำ

เนื่องจากการเติบโตขึ้นอย่างรวดเร็วของอุตสาหกรรมเพาะเลี้ยงนกกระทาในประเทศไทยทำให้เกษตรกรไทยให้ความสนใจกับอาชีพนี้เป็นอย่างมาก แต่อุปสรรคที่พบคือ เกษตรกรจะนิยมเลี้ยงเฉพาะนกกระทาเพศเมียมากกว่าเพศผู้ เพราะเพศเมียจะให้ผลผลิตที่สูงกว่าเพศผู้ เพศผู้จะถูกขายทิ้งเพื่อบริโภคเนื้อ แต่ปัญหาที่เกิดขึ้นคือเกษตรกรไม่สามารถรู้เพศของนกกระทาที่ฟักออกมาจากไข่ได้ในทันทีที่ฟัก จะต้องใช้วิธีสังเกตความแตกต่างของขนหน้าอกเมื่อลูกนกกระทาอายุได้ 21 วัน หลังฟักออกจากไข่ ซึ่งทำให้ต้นทุนการผลิตเพิ่มสูงขึ้น โดยที่หากเราสามารถระบุหรือทำนายเพศของนกกระทาว่าเป็นเพศผู้หรือเพศเมียได้ตั้งแต่ก่อนทำการฟัก จะทำให้เกษตรกรไทยลดต้นทุนการเลี้ยงให้ต่ำลง

ในปัจจุบันวิธีที่เกษตรกรไทยใช้สำหรับการแยกเพศของนกกระทาก่อนทำการฟักแบบที่สืบทอดกันมาคือ ใช้การสังเกตรูปทรงไข่ ที่เรียวยาวแหลมจะฟักออกมาเป็นเพศผู้ ส่วนรูปทรงไข่ที่ค่อนข้างกลมจะฟักออกมาเป็นเพศเมีย ซึ่งในงานวิจัยการแยกเพศนกกระทาจากการพิจารณาลักษณะภายนอกของไข่ โดยพงษ์พัฒน์ สิงห์ศรี จักริน สุขสวัสดิ์ชน และ อูรีรัฐสุขสวัสดิ์ชน [1] ได้ทำการทดลองและนำเสนอค่าความถูกต้องในการจำแนกเพศพบว่า การใช้ปัจจัย เพียง 2 ปัจจัยตามที่สืบทอดกันมาของเกษตรกรไทย มีค่าความถูกต้องในการคัดแยกเพศของนกกระทาเพียง 50% เท่านั้น และเมื่อเพิ่มปัจจัยในการจำแนกเพศเป็น 7 ปัจจัย พบว่าผลการทดลองโดยวิธีการแยกประเภทแบบต้นไม้ตัดสินใจให้ความถูกต้องจากการจำแนกเพศนกกระทาสูงถึง 77.46%

จากปัญหาวิจัยและผลการทดลองเบื้องต้นที่ได้นำเสนอในบทความวิจัย [1] แล้วนั้น ผู้วิจัยพบว่ายังคงมีปัญหาคือจะต้องทำการศึกษาวิจัยเพิ่มเติมคือ “ปัจจัยจากลักษณะภายนอกของไข่นกกระทาทั้ง 7 ปัจจัย ปัจจัยใดเป็นปัจจัยที่สัมพันธ์กันและเพียงพอต่อการจำแนกเพศของนกกระทาได้ดีที่สุด” และ “วิธีการแยกประเภทแบบใดเหมาะสมที่สุดในการจำแนกเพศนกกระทา”

ดังนั้นในงานวิจัยนี้ผู้วิจัยได้นำเสนอแบบจำลองการทำนายเพศของนกกระทาโดยการประยุกต์ใช้เทคนิควิธีการแยกประเภทร่วมกับปัจจัยภายนอกของไข่นกกระทา ซึ่งจะทำการ

เก็บรวบรวมข้อมูลของปัจจัยที่มีผลต่อการทำนายเพศตั้งแต่ก่อนทำการฟักจนถึงคัดแยกเพศได้หลังจากฟัก และทำการทดลองจะใช้การคัดเลือกความสัมพันธ์ของปัจจัยสองวิธีการคือ (1) วิธีการ Best First Search (2) วิธีการวิเคราะห์การถดถอยพหุคูณแบบขั้นตอน (Stepwise Multiple Regression Analysis) จากนั้นใช้ข้อมูลปัจจัยเด่นที่มีความสัมพันธ์กันมาสร้างแบบจำลองด้วยเทคนิคการแยกประเภท (Classification techniques) เช่น J48 ADTree REPTree LibSVM SMO Multilayer Perceptron RBFnetwork Naïve Bayes NNge และ Conjunctive Rule เป็นต้น เพื่อทำการทดลองเปรียบเทียบความถูกต้องในการทำนายเพศของนกกระทาต่อไป

ผลการทดลองจากวิจัยนี้จะสามารถพัฒนาเป็น โปรแกรมประยุกต์เพื่อใช้สำหรับทำนายเพศของนกกระทา ซึ่งจะเป็ประโยชน์ต่อเกษตรกรที่เลี้ยงนกกระทาในการลดต้นทุนการค่าใช้จ่ายของเกษตรกรต่อไปในอนาคต

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 งานวิจัยที่ศึกษาเกี่ยวกับนกกระทาและไข่

ในด้านรูปทรงไข่ ได้มีโครงการงานวิจัยจากโรงเรียนสมเด็จพระปิยะดาม [2] ได้ทำการศึกษาว่ารูปทรงของไข่จะมีผลกระทบต่อเพศของไก่หรือไม่ โดยทำการทดลองแบ่งรูปทรงของไข่เป็นรูปทรงรี และรูปทรงค่อนข้างกลม ผลการศึกษาวิจัยพบว่าไก่เนื้อ และไก่ไข่ที่มีรูปทรงรีจะมีอัตราส่วนเพศผู้สูงกว่าเพศเมีย ส่วนไข่ที่มีรูปทรงค่อนข้างกลมจะมีอัตราส่วนเพศเมียสูงกว่าเพศผู้

ในด้านการใช้ดัชนีรูปทรงไข่ (Shape index) จากงานวิจัยของ Maclaury และคณะ [3] ได้ศึกษาวิจัยโดยใช้สูตรการหาดัชนีรูปทรงไข่ ($W/L * 100$) ร่วมกับวิธีการหาปริมาตรของไข่เพื่อศึกษาโอกาสการฟักออกของไข่ที่มีเชื้อของนกกระทาพบว่าค่าสหสัมพันธ์ทางลบระหว่างน้ำหนักไข่ และการฟักออกของไข่นกกระทาญี่ปุ่น ดัชนีรูปทรงไข่ เป็นลักษณะภายนอกอย่างหนึ่ง โดยทำการแบ่งกลุ่มสำหรับทดลองออกเป็น 4 กลุ่มพบว่า ในไข่ที่มีรูปทรงปกติจะมีอัตราการฟักออกที่สูงขึ้นเพราะเนื่องจากการมีพื้นที่ภายในที่เหมาะสมสำหรับตำแหน่งของเอ็มบริโอ (Embryo) ในไข่ ซึ่งเป็นผลทำให้ไข่ที่มีรูปทรงกลมหรือรียาวจะมีอัตราการฟักออกที่ต่ำกว่าไข่รูปทรงปกติ งานวิจัย

[4] เป็นการหาประสิทธิภาพในการฟักรอกจากความสัมพันธ์ระหว่างความหนาของเปลือกไข่ ความพรุนของเปลือกไข่ คำนวณรูปทรงไข่ และการสูญเสียน้ำหนักของไข่ จากการฟัก ผลการทดลองพบว่ามีความแตกต่างทางสถิติอย่างมีนัยสำคัญที่ $P < 0.001$

งานวิจัย [5, 6] เป็นงานวิจัยที่มุ่งเน้นหาวิธีการเพิ่มผลผลิตของเนื้อและไข่นกกระทา ส่วนงานวิจัยต่างประเทศยังไม่มีงานวิจัยใดที่ศึกษาและวิจัยการจำแนกหรือแยกเพศนกกระทาจากการพิจารณาลักษณะภายนอกของไข่ ส่วนใหญ่เป็นงานวิจัยที่เกี่ยวกับปัจจัยที่มีผลต่อการฟักออกและคุณภาพของไข่ การเพิ่มผลผลิตเป็นต้น เช่น งานวิจัยการพยากรณ์ น้ำหนักลูกนกกระทาแรกเกิด [7] โดยวิธีพยากรณ์ น้ำหนักของลูกนกกระทาแรกเกิดจะช่วยให้ policy makers สามารถตัดสินใจก่อนที่จะนำไข่เข้าทำการฟักเนื่องจาก ถ้าน้ำหนักลูกนกกระทามากหรือน้อยกว่าปกติ ลูกนกจะอ่อนแอ และโอกาสการรอดหลังเกิดจนอายุได้ 4 สัปดาห์จะมีน้อย

2.2 ต้นไม้ตัดสินใจ

การสร้างต้นไม้ในการตัดสินใจ [8] โดยปกติจะอยู่ในรูปของกฎ “ IF (conditions) Then (results) ” ต้นไม้ในการตัดสินใจประกอบไปด้วย (1) โหนดเพื่อใช้แทนคุณลักษณะ (2) กิ่งแทนเงื่อนไขในการทดสอบ (3) ใบแทนผลลัพธ์หรือคลาสคำตอบ

ซึ่งในการเลือกคุณลักษณะมาเป็นโหนดในต้นไม้จะใช้ค่าเกน (Information Gain) โดยคุณลักษณะที่มีค่าเกนสูงที่สุดจะถูกเลือกมาเป็นโหนดราก ค่าเกนนี้คำนวณได้จากค่าเอนโทรปี (Entropy) ทั้งหมดของชุดข้อมูล

2.3 ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine หรือ SVM) [9]

แนวคิดของทฤษฎีนี้ใช้เพื่อหาระนาบในการตัดสินใจที่เหมาะสมที่สุดในการแบ่งข้อมูลออกเป็นสองส่วน โดยพยายามสร้างเส้นตรงหรือระนาบหลายมิติ (Hyper plane) ที่เหมาะสมตรงกึ่งกลางระหว่างกลุ่มให้ระยะห่างระหว่างขอบเขตของทั้งสองกลุ่มมากที่สุด SVM มีความสามารถในการแยกข้อมูลที่ไม่เป็นเชิงเส้นได้

2.4 โครงข่ายประสาทเทียม

- โครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ (Back-propagation Neural Network) [8] เป็นขั้นตอนที่ใช้ในการเรียนรู้ของโครงข่ายประสาทวิธีหนึ่งที่นิยมใช้ใน Multilayer perceptron เพื่อปรับค่าน้ำหนักในเส้นเชื่อมต่อระหว่างนิวรอนแต่ละโหนดให้เหมาะสม โครงข่ายประสาทเทียมประกอบด้วยชั้นของ node 3 ชั้น ได้แก่ ชั้นข้อมูลนำเข้า (Input layer) ชั้นซ่อน (Hidden layer) และชั้นผลลัพธ์ (Output layer)

- ข่ายงานเรเดียลเบสิคฟังก์ชัน (Radial Basis Function network) [10] จัดเป็นข่ายงานที่มีสถาปัตยกรรมแบบไปข้างหน้าและมีหลายชั้น ข่ายงานประสาทเทียมแบบหลายชั้นมีความสามารถในการทำ non-linear mapping และเมื่อปัญหานั้นมีความยากเพิ่มขึ้น จะใช้วิธีเพิ่มชั้นของเพอร์เซปตรอนไปเรื่อย ๆ ข้อจำกัดของการใช้เส้นแบ่งแยก คือ ไม่อาจจำกัดขอบเขตของกลุ่มของข้อมูลตามธรรมชาติของข้อมูลได้

2.5 ทฤษฎีของเบย์ [8]

เป็นทฤษฎีที่ใช้ทฤษฎีความน่าจะเป็นมาช่วยในการเรียนรู้เพื่อสร้างตัวแบบให้อยู่ในรูปความน่าจะเป็น ข้อดีคือ สามารถใช้ข้อมูลและความรู้ก่อนหน้ามาช่วยเรียนรู้ได้ ซึ่งจะทำให้ประสิทธิภาพในการเรียนรู้ดีเทียบเท่ากับวิธีการเรียนรู้ประเภทอื่น

2.6 ขั้นตอนวิธีการเพื่อนบ้านที่ใกล้ที่สุด (Nearest Neighbor) [11]

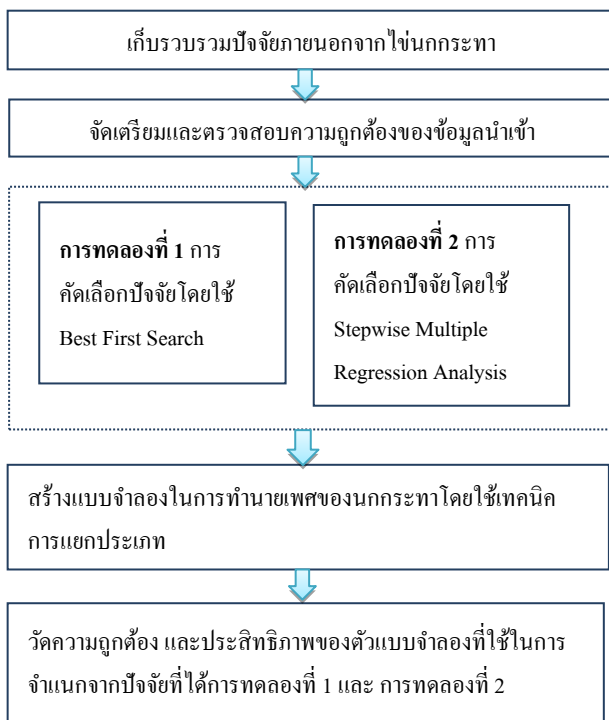
เป็นการหาระยะห่างระหว่างแต่ละตัวแปรในข้อมูล ซึ่งวิธีนี้จะเหมาะสำหรับข้อมูลแบบตัวเลข แต่ตัวแปรที่เป็นค่าแบบไม่ต่อเนื่องนั้นก็ยังสามารถทำได้ จากนั้นทำการรวมค่าระยะห่างของ Attribute ทุกค่าที่วัดได้ เมื่อสามารถคำนวณระยะห่างระหว่างเงื่อนไขหรือกรณีต่าง ๆ ได้ จากนั้นก็เลือกชุดของเงื่อนไขที่ใช้จัดคลาสมาเป็นฐานสำหรับการจัดคลาสในเงื่อนไขใหม่ ๆ ทำให้ตัดสินใจขอบเขตของจุดข้างเคียงที่ควรเป็นได้

2.7 กฎการเชื่อมต่อกัน (Conjunctive rule) [12]

เป็นกฎที่ใช้ลักษณะของการให้เหตุผลแบบอุปนัย ซึ่งเป็นวิธีการสรุปผลมาจากการค้นหาความจริงจากการสังเกต หรือการทดลองหลายครั้งจากกรณีย่อย ๆ แล้วนำมาสรุปเป็นความรู้แบบทั่วไป

3. วิธีการนำเสนอ

3.1 ขั้นตอนการดำเนินงานวิจัย



ภาพที่ 1 : ขั้นตอนการดำเนินการวิจัย

3.2 การเก็บรวบรวมข้อมูล

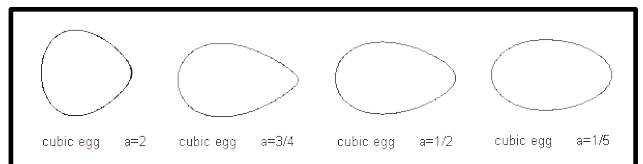
ไข่นกกระทาที่ใช้ในการทำทดลอง ได้จัดซื้อจากฟาร์มนกกระทา จังหวัดสกลนคร จำนวนทั้งสิ้น 1,000 ฟอง ทำการทดลองในช่วงเดือน ตุลาคม ถึง มกราคม ซึ่งเป็นช่วงที่มีอัตราการฟักเป็นตัวสูงที่สุด ช่วงเริ่มต้นการทดลองไข่นกกระทาแต่ละฟองจะถูกเก็บบันทึกข้อมูลที่เป็นปัจจัยภายนอกซึ่งประกอบไปด้วย 7 ปัจจัย คือ ความสูง ความกว้าง น้ำหนัก รูปทรง ดัชนี รูปทรงไข่ ไข่ของเปลือกไข่ และลักษณะด้านแหลมของไข่นกกระทา จนกระทั่งไข่ฟักออกมาเป็นตัว และคัดแยกเพศได้ ซึ่งข้อมูลที่เก็บได้จะเหลือเพียง 661 ชุดข้อมูล เนื่องจากไข่บางฟองไม่ฟักออกเป็นตัว และนกกระทาบางตัวไม่สามารถกำหนดเพศได้ว่าเป็นเพศผู้หรือเพศเมีย (เกษตรกรจะเรียกนกชนิดนี้ว่า นกกระทา) ตัวอย่างรูปแบบการเก็บข้อมูลแสดงในตารางที่ 1

ตารางที่ 1: ปัจจัยภายนอก 7 ปัจจัยของไข่นกกระทาและรูปแบบการจัดเก็บข้อมูลนำเข้า

ข้อมูล	รายละเอียดข้อมูล
ความกว้าง (Width)	ค่าจริง
ความสูง (Length)	ค่าจริง
น้ำหนัก (Weight)	ค่าจริง
ค่าดัชนีรูปทรงไข่ (SI) (SI) = ((W/L)x100)	ค่าจริง
ลายไข่ (EC) (ภาพที่ 2)	0 = ลายป้าน, ลายเต็ม 1 = ลายดาว
รูปทรงไข่ (Shape) (ภาพที่ 3)	1 = เรียว ; 2=ค่อนข้างเรียว 3 = ค่อนข้างกลม; 4 = กลม
ลักษณะด้านแหลม (SE)	0 = รูปทรงมน 1 = รูปทรงแหลม



ภาพที่ 2 : ตัวอย่างลายไข่ (ซ้าย) ไข่ลายดาว (ขวา) ไข่ลายเต็ม



ภาพที่ 3 : รูปทรงไข่กลม ค่อนข้างกลม ค่อนข้างเรียว และเรียว เรียงจากซ้ายไปขวา

3.3 การหาความสัมพันธ์ของปัจจัยในการจำแนกเพศนกกระทา

งานวิจัยนี้ได้ทำการหาความสัมพันธ์ของปัจจัยที่ใช้ในการทำนายเพศของนกกระทาด้วยวิธีการ Best First Search และ Stepwise Multiple Regression Analysis

3.3.1 จากการทดลองหาความสัมพันธ์ของปัจจัยด้วยวิธีการ Best First Search โดยใช้โปรแกรม WEKA รุ่น 3.7.8 พบว่ามีปัจจัยที่มีความสัมพันธ์กันจำนวน 2 ปัจจัย คือ ความกว้าง และลักษณะด้านแหลมของไข่

3.3.2 จากการทดลองหาความสัมพันธ์ของปัจจัยด้วยวิธีการ Stepwise Multiple Regression Analysis โดยใช้โปรแกรม SPSS พบว่ามีปัจจัยที่มีความสัมพันธ์กันจำนวน 2 ปัจจัย คือน้ำหนัก และลักษณะด้านแหลมของไข่

3.3 การสร้างตัวแบบในการทำนายเพศนกกระทาด้วยเทคนิคการแยกประเภท

ตัวแบบที่ใช้ในการทดสอบการจำแนกเพศของนกกระทาประกอบด้วย 10 ตัวแบบดังนี้

- Decision Trees: (1) J48 (2) ADTree (3) REPTree
- Support Vector Machine: (4) LibSVM (5) SMO
- ANN: (6) MultilayerPerceptron (7) RBFnetwork
- (8) Naïve Bayes
- Nearest Neighbor: (9) NNge
- Rule Based: (10) Conjunctive Rule

ในงานวิจัยนี้ใช้โปรแกรมประยุกต์ WEKA 3.7.8 ในการทำการทดลอง และกำหนดฟังก์ชันที่ใช้ในการเรียนรู้เป็น 10-Fold Cross Validation โดยได้ทำการปรับค่าพารามิเตอร์ของตัวแบบดังแสดงในตารางที่ 2

ตารางที่ 2: ค่าพารามิเตอร์ที่ใช้ในแต่ละตัวแบบ

ตัวแบบ	พารามิเตอร์ที่ปรับค่า	ค่าที่ปรับ
J48	minimum number object	2 - 50
ADTree	Boosting iterations	10 – 30
REPTree	minimum number of instances per leaf	2 - 20
LibSVM	cost	1 – 50
SMO	gamma	0.0 0.2 0.4 0.6 0.8 1.0
MultilayerPerceptron	Hidden Layer Learning rate Training time	a, t, i และ o 0.2 และ 0.3 200, 300, 400, 500
RBFnetwork	number of clusters	1 – 20
NNge	number of attempts number of folder for mutual information	5 – 10 5 -10
conjunctiveRule	fold	3 - 15

4. การทดสอบและประเมินผล

งานวิจัยนี้ได้มีการวัดประสิทธิภาพและเปรียบเทียบผลการทำนายการจำแนกเพศของนกกระทาจากปัจจัยภายนอกของไข่นกกระทาโดยใช้ตัวแบบการแยกประเภท โดยใช้ค่าจากตาราง Confusion Matrix ตามตัวอย่างในตารางที่ 3 เพื่อนำมาหาค่า Accuracy, ค่า Precision, ค่า Recall และค่า F-Measure [10] ได้จากสมการดังต่อไปนี้

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F - Measure = \frac{2 x (Recall x Precision)}{Recall + Precision} \tag{4}$$

ตารางที่ 3: Confusion matrix สำหรับการวัดประสิทธิภาพ

Actual Class	Predicted Class	
	เพศผู้	เพศเมีย
เพศผู้	TN	FP
เพศเมีย	FN	TP

ตารางที่ 4: ผลการวัดประสิทธิภาพความถูกต้องจากตัวแบบเมื่อใช้ปัจจัยที่ได้จากวิธีการ Best First Search

Algorithms	Precision (%)	Recall (%)	F-Measure (%)	Accuracy (%)
J48	68.45	80.11	73.82	69.74
ADTree	69.59	76.7	72.97	69.74
REPTree	67.97	74.15	70.92	67.62
LibSVM	69.39	77.27	73.12	69.74
SMO	70.14	70.74	70.44	68.38
MultilayerPerceptron	68.91	78.69	73.47	69.74
RBFnetwork	68.94	77.56	72.99	69.44
naïveBayes	69.41	76.7	72.87	69.59
NNge	62.32	62.5	62.41	59.91
conjunctiveRule	70.06	70.45	70.25	68.23

ตารางที่ 5: ผลการวัดประสิทธิภาพความถูกต้องจากตัวแบบเมื่อใช้

ปัจจัยที่ได้จากวิธีการ Stepwise Multiple Regression Analysis

Algorithms	Precision (%)	Recall (%)	F-Measure (%)	Accuracy (%)
J48	68.59	74.43	71.39	68.23
ADTree	67.71	73.86	70.65	67.32
REPTree	68.36	72.44	70.34	67.47
LibSVM	70.42	76.42	73.30	70.35
SMO	70.14	70.74	70.44	68.38
MultilayerPerceptron	69.19	75.28	72.11	68.99
RBFnetwork	69.09	75.57	72.18	68.99
naïveBayes	69.33	76.42	72.70	69.44
NNge	58.66	59.66	59.15	56.13
conjunctiveRule	71.18	70.17	70.67	68.99

จากผลการทดลองที่ได้ตามตารางที่ 4 และ 5 พบว่าเมื่อใช้ปัจจัยที่ได้จากวิธีการ Best First Search ตัวแบบที่ให้ผลความถูกต้องสูงสุดคือ ADTree ให้ค่าความถูกต้อง = 69.74% ค่าความแม่นยำ = 69.59% ค่า Recall = 76.70 และค่าความถ่วงดุล (F-Measure) = 72.97% แต่ตัวแบบทั้ง 9 ตัวแบบ (ยกเว้น NNge) จะให้ค่าความถูกต้องที่ค่อนข้างใกล้เคียงกันประมาณ 67.62 - 69.74 % และเมื่อใช้ปัจจัยที่ได้จากวิธีการ Stepwise Multiple Regression ตัวแบบที่ให้ผลความถูกต้องสูงสุดคือ LibSVM ให้ค่าความถูกต้อง = 70.35% ค่าความแม่นยำ = 70.42% และค่า Recall = 76.42% และค่าความถ่วงดุล (F-Measure) = 73.30%

6. สรุปผลการวิจัย

จากแบบจำลองการทำนายเพศของนกกระทาจากคุณลักษณะภายนอกของไข่โดยใช้เทคนิคการแยกประเภทที่ได้นำเสนอในงานวิจัยนี้ ซึ่งได้ทำการวิเคราะห์หาปัจจัยที่มีความสัมพันธ์กัน ผลการวิจัยพบว่าวิธีการ Best First Search มี 2 ปัจจัยที่มีความสัมพันธ์กันในการทำนายเพศของนกกระทาได้แก่ ความกว้าง และลักษณะด้านแหลมของไข่ และเมื่อนำปัจจัยที่ได้มาทำการทดลองในตัวแบบการจำแนกเพศพบว่า ตัวแบบ ADTree ให้ผลความถูกต้องมากที่สุด แต่ตัวแบบอื่นๆ (ยกเว้น NNge) ก็ให้ผลความถูกต้องที่ค่อนข้างใกล้เคียงกันแต่

สำหรับปัจจัยที่มีความสัมพันธ์กันที่ได้จากวิธีการ Stepwise Multiple Regression Analysis มีจำนวน 2 ปัจจัยเช่นกัน แต่เป็นน้ำหนัก และลักษณะด้านแหลมของไข่ ผลการวัดประสิทธิภาพจากตัวแบบพบว่า LibSVM ให้ผลความถูกต้องที่สูงกว่าตัวแบบอื่น และมากกว่าตัวแบบ ADTree ที่ใช้ปัจจัยความกว้าง และลักษณะด้านแหลมของไข่ในการทำนาย

ดังนั้นผลการดำเนินการวิจัยที่ได้นำเสนอมานั้น ผู้วิจัยพบว่าจากปัจจัยภายนอกของไข่ในนกกระทาทั้ง 7 ปัจจัยที่ใช้อยู่สำหรับการทำนายเพศนกกระทานั้น สามารถลดปัจจัยลงให้เหลือเพียง 2 ปัจจัยที่มีความสัมพันธ์กัน ซึ่งโมเดลของตัวแบบที่สามารถนำไปพัฒนาต่อไปได้คือ LibSVM ร่วมกับปัจจัยน้ำหนัก และลักษณะด้านแหลมของไข่ ที่ให้ค่าความถูกต้องสูงถึง 70% แต่ยังมีค่าความถูกต้องน้อยกว่าการใช้ทั้ง 7 ปัจจัย ซึ่งมีค่าความถูกต้องถึง 77%

อย่างไรก็ตามผู้วิจัยยังคงต้องศึกษาและทดลองเพิ่มเติมเพื่อหาโมเดลที่เหมาะสมยิ่งขึ้นเพื่อใช้ในการแยกเพศ ซึ่งจะช่วยให้ผลความถูกต้องในการแยกเพศมีความถูกต้องเพิ่มขึ้น อีกทั้งวิเคราะห์ปัจจัยภายนอกจากไข่ในนกกระทาที่มีผลต่อการทำนายเพศนกกระทาเป็นเพียงเครื่องมือในการประเมินเพศเบื้องต้นเท่านั้น แต่ปัจจัยที่ทำให้เกิดเพศแท้จริงแล้วอาจมีปัจจัยหรือตัวแปรที่มากกว่าตัวแบบที่สร้างขึ้น การวิจัยหาปัจจัยอื่นจึงเป็นเรื่องสำคัญในการหาปัจจัยที่จะสามารถกำหนดเพศนกกระทาในระยวต่อไป ซึ่งถือว่าเป็นประโยชน์อย่างยิ่งในการนำไปพัฒนาต่อเพื่อหาปัจจัยที่มีความแน่นอนมากในการประเมินเพศนกกระทาล่วงหน้า ซึ่งทำให้เกษตรกรผู้เลี้ยงนกกระทามีต้นทุนการเลี้ยงที่ต่ำ และพัฒนาอุตสาหกรรมการเพาะเลี้ยงนกกระทาของประเทศไทยต่อไป

7. กิตติกรรมประกาศ

งานวิจัยนี้ได้รับทุนสนับสนุนจากเงินอุดหนุนทุนวิจัยคณะวิทยาการสารสนเทศ มหาวิทยาลัยบูรพา ประจำปีงบประมาณ 2556 และขอขอบคุณคณะเกษตรและเทคโนโลยีนครพนม มหาวิทยาลัยนครพนม ที่เอื้อเฟื้อสถานที่ในการฝึกเลี้ยงและเก็บข้อมูลนกกระทา

เอกสารอ้างอิง

- [1] พงศ์พัฒน์ สิงห์ศรี จักริน สุขสวัสดิ์ดิษฐ์ และ อุรีรัฐ สุขสวัสดิ์ดิษฐ์. 2556. การประยุกต์ใช้เทคนิคการแยกประเภทในการวิเคราะห์ปัจจัยที่มีผลต่อการจำแนกเพศของนกกกระทาที่พิจารณาจากคุณลักษณะภายนอกของไข่ นกกกระทา. NCCIT, 9, pp. 240-247, 2013.
- [2] โรงเรียนสมเด็จพระพิทยาคม. โครงการหารูปทรงของไข่มีความสัมพันธ์ต่อเพศของไก่. วันที่ค้นข้อมูล 18 กันยายน 2555, เข้าถึงได้จาก <http://www.somdetpit.ac.th/P-shap.htm>
- [3] Maclaury, D.W., W.M. Insko, Jr., J.J. Begin and T.H. Johnson. (1973). Shape index versus hatchability of fertile eggs of Japanese quail (*Coturnix coturnix japonica*). *Poultry Sci*, 52, 558-562.
- [4] Mehmet Kenan TÜRKÜYLMAZ, Evrim DERELI, Tayfun ŞAHİN. "Effects of Shell Thickness, Shell Porosity, Shape Index and Egg Weight Loss on Hatchability in Japanese Quail (*Coturnix coturnix japonica*)". *Kafkas Üniv. Vet. Fak. Derg.*, 11(2), pp. 147-150, 2005.
- [5] วราภรณ์ ถิ่นประวดี นงเยาว์ จันทราช และสุเมธี กิตติพงษ์ไพศาล, "ปัญหาพิเศษเรื่องสมรรถภาพการผลิตนกกกระทาเนื้อที่ได้รับอาหารเสริมมันสำปะหลังที่หมักด้วย E.M. (Effective Microorganism) ในระดับต่างกัน". วิทยาศาสตร์บัณฑิต, สำนักวิทยบริการและเทคโนโลยีสารสนเทศ, มหาวิทยาลัยราชภัฏจันทรเกษม, 2547.
- [6] มณี โททอง, นงเยาว์ จันทราช และสุเมธี กิตติพงษ์ไพศาล, "สมรรถนะการผลิตของนกกกระทาเนื้อที่ได้รับการเสริมสมุนไพรฟ้าทะลายโจรในอาหาร". วิทยาศาสตร์บัณฑิต, สำนักวิทยบริการและเทคโนโลยีสารสนเทศ, มหาวิทยาลัยราชภัฏจันทรเกษม, 2547.
- [7] Khurshid, A., M. Farooq, F. R. Durrani, K. Sarbiland and N. Chand. "Predicting Egg Weight, Shell Weight, Shell Thickness and Hatching Chick Weight of Japanese Quails Using Various Egg Traits as Regressors". *Asian Network for Scientific Information*, 2005
- [8] บุญเสริม กิจศิริกุล. 2546. ปัญญาประดิษฐ์. กรุงเทพฯ ; ภาควิชาวิศวกรรมคอมพิวเตอร์ จุฬาลงกรณ์มหาวิทยาลัย.
- [9] เดช ชรรณศิริ และพยุ่ง มีสัจ. 2544. การจำแนกข้อมูลด้วยวิธีแบบร่วมตัดสินใจจากพื้นฐานของเทคนิคต้นไม้ตัดสินใจ เทคนิคโครงข่ายประสาทเทียม และเทคนิคซัพพอร์ตเวกเตอร์แมชชีน ร่วมกับการเลือกตัวแทนที่เหมาะสมด้วยขั้นตอนวิธีเชิงพันธุกรรม. วารสารวิชาการพระจอมเกล้าพระนครเหนือ, 21(2):295-296.
- [10] สิริภัทร เชี่ยวชาญวัฒนา . 2552. เอกสารประกอบการสอนวิชา ายงานประสาทเทียม (Artificial Neural Networks).
- [11] Brent Martin. "INSTANCE-BASED LEARNING: Nearest Neighbour with Generalisation". 1995.
- [12] Mohd Fauzi bin Othman, Thomas Moh Shan Yau . "Comparison of Different Classification Techniques Using WEKA for Breast Cancer", *Biomed 06, IFMBE Proceedings 15*, pp. 520-523, 2007.

การให้ค่าน้ำหนักบนคุณลักษณะที่ใช้ในการสืบค้นข้อมูลของผู้หลอกลวงขายสินค้าออนไลน์

Weighted Features for Online Shopping Fraud Detection

นิธิภัทร์ เพชรมงคลจรัส (Nitipat Petmongkonjaras)¹ และ ธนัสนี เพียรตระกูล (Tanasanee Phienthrakul)²

ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยมหิดล

¹nitipat_boo@hotmail.com, ²tanasanee.ph@gmail.com

บทคัดย่อ

บทความนี้นำเสนอวิธีการในการกำหนดค่าน้ำหนักให้กับแต่ละคุณลักษณะเพื่อใช้ในการสืบค้นข้อมูลของผู้กระทำการหลอกลวงขายสินค้าบนเครือข่ายอินเทอร์เน็ต วิธีการรีลีฟถูกนำมาใช้เพื่อช่วยในการประเมินความสำคัญของแต่ละคุณลักษณะ ค่าน้ำหนักที่ได้จากวิธีการรีลีฟจะถูกปรับปรุงและนำมาใช้เปรียบเทียบข้อมูลของผู้ต้องสงสัย กับข้อมูลของผู้เคยทำการหลอกลวงที่ได้มีการเก็บรวบรวมไว้ ผลการทดสอบแสดงให้เห็นว่าวิธีการที่นำเสนอสามารถช่วยให้การสืบค้นข้อมูลของผู้ต้องสงสัย และให้ผลการสืบค้นที่มีความแม่นยำมากขึ้นเมื่อเปรียบเทียบกับวิธีการสืบค้น โดยไม่ได้กำหนดค่าน้ำหนัก

คำสำคัญ: การหลอกลวงซื้อขายสินค้าออนไลน์ การกำหนดค่าน้ำหนักในการสืบค้น วิธีการรีลีฟ

Abstract

This paper proposed a method for estimating the weight of each attribute in an online shopping fraud detection system. Relief algorithm was applied to evaluate each attribute. The weights from Relief were adjusted and these weights were used for comparing between the incomplete data of suspects and the data from a criminal database. The experimental results showed that the proposed method can search the data of suspects. The average percentages of similarity were reported and compare to the unweighted search.

Keyword: online shopping fraud, weighted search, Relief algorithm.

1. บทนำ

การซื้อขายแลกเปลี่ยนสินค้าบนเครือข่ายอินเทอร์เน็ตได้รับความนิยมเพิ่มมากขึ้น เนื่องจากเป็นช่องทางการติดต่อซื้อขายที่ลงทุนไม่มากนัก และสามารถเข้าถึงผู้คนได้เป็นจำนวนมาก ผู้ขายบางรายไม่จำเป็นต้องมีหน้าร้านจริง หรือไม่จำเป็นต้องกักตุนสินค้าไว้เป็นจำนวนมาก ทำให้สามารถกำหนดราคาได้ต่ำกว่าร้านค้าที่มีหน้าร้านจริง และยังสามารถเสนอขายได้ทั้งขายปลีกและขายส่ง ด้านผู้ซื้อเองก็สามารถศึกษาข้อมูล และเปรียบเทียบข้อมูลก่อนการตัดสินใจได้อย่างสะดวก ทำให้เกิดการแลกเปลี่ยนซื้อขายสินค้าผ่านทางเว็บไซต์ต่าง ๆ กันมาก

แม้ว่าการซื้อขายสินค้าผ่านทางเครือข่ายอินเทอร์เน็ตจะสะดวก แต่เป็นการตกลงซื้อขายโดยผู้ซื้อและผู้ขายไม่จำเป็นต้องรู้จักหรือเห็นหน้ากัน รวมถึงไม่ได้เห็นสินค้าจริงก่อนการซื้อ จึงเป็นช่องทางให้มิจฉาชีพกระทำการหลอกลวงผู้ซื้อ โดยหลอกให้โอนเงินผ่านธนาคารแต่ไม่ส่งมอบสินค้าให้ หรือส่งสินค้าที่มีตำหนิ สินค้าปลอม หรือ สินค้าไม่ตรงกับในประกาศ ก่อให้เกิดความเสียหายแก่ผู้ถูกหลอกลวง และยากต่อการติดตามตัวผู้กระทำความผิดมาลงโทษ เนื่องจากผู้หลอกลวงมักไม่ได้ให้ข้อมูลจริงทั้งหมด และยังสามารถเปลี่ยนแปลงข้อมูลหลังจากหลอกลวงแล้ว อาทิเช่น นามแฝง ชื่อ นามสกุล ที่อยู่ หมายเลขบัญชี และ เบอร์โทรศัพท์ ทำให้มีผู้ถูกหลอกลวงซ้ำเกิดขึ้น

ผู้ซื้อสินค้าสามารถป้องกันตัวเองจากการถูกหลอกลวงได้ โดยการสืบค้นและตรวจสอบประวัติของผู้ขายสินค้า ผ่านทางหน้าเว็บไซต์ที่ประกาศขายสินค้า หรือ สืบค้นจากเว็บไซต์ที่เคยมีผู้ถูกหลอกลวงได้ให้ข้อมูลไว้ แต่ยังไม่มียังไม่มีเว็บไซต์ใดที่ทำการรวบรวมข้อมูลทั้งหมดไว้ในที่เดียว และการที่ผู้หลอกลวง

สามารถเปลี่ยนแปลงข้อมูลส่วนบุคคลของตนเอง ทำให้อาจไม่สามารถสืบค้นพบได้โดยง่าย

งานวิจัยนี้จึงได้เสนอแนวทางในการสืบค้นข้อมูลของผู้ต้องสงสัยที่จะทำการหลอกลวงขายสินค้าผ่านทางเครือข่ายอินเทอร์เน็ต โดยได้ทำการออกแบบและพัฒนาระบบฐานข้อมูลสำหรับจัดเก็บข้อมูลของผู้กระทำการหลอกลวงขายสินค้าบนเครือข่ายอินเทอร์เน็ต และออกแบบระบบการสืบค้นข้อมูลของผู้ต้องสงสัย เปรียบเทียบกับข้อมูลที่มีอยู่ในฐานข้อมูล โดยใช้การกำหนดค่าถ่วงน้ำหนักที่ได้จากเรียนรู้ด้วยวิธีการลิฟต์บนข้อมูลที่มีอยู่ และสมมติฐานเพิ่มเติมที่ได้กำหนดขึ้น

2. ทบทวนวรรณกรรม

ในบทความนี้เลือกที่จะใช้วิธีลิฟต์เพื่อช่วยในการกำหนดค่าถ่วงน้ำหนักสำหรับแต่ละคุณลักษณะที่ใช้ในการสืบค้นข้อมูล ซึ่งโดยทั่วไปแล้ววิธีลิฟต์จะนำมาใช้ในการลดจำนวนคุณลักษณะของข้อมูล แต่แนวคิดของงานวิจัยนี้ยังคงเลือกที่จะเก็บคุณลักษณะทั้งหมดที่ไว้ใช้ในการสืบค้น แต่ให้ความสำคัญกับแต่ละคุณลักษณะไม่เท่ากัน ในการทดสอบความสามารถในการสืบค้นข้อมูล ซัพพอร์ตเวกเตอร์แมชชีนถูกเลือกนำมาใช้เพื่อแสดงให้เห็นว่าการให้ค่าน้ำหนักของคุณลักษณะด้วยวิธีการดังกล่าวมีความเหมาะสมมากน้อยเพียงใด สำหรับทฤษฎีที่เกี่ยวข้องกับวิธีลิฟต์และซัพพอร์ตเวกเตอร์แมชชีนจะถูกกล่าวโดยสรุปดังต่อไปนี้

2.1 การคัดเลือกคุณลักษณะแบบวิธีลิฟต์

การเลือกคุณลักษณะของข้อมูลแบบวิธีลิฟต์เป็นอัลกอริทึมที่เสนอโดย Kira and Rendell [1] ซึ่งได้รับความนิยมเป็นอย่างมาก เนื่องจากสามารถเข้าใจและแปลความหมายได้ง่าย ทำงานได้เร็ว มีประสิทธิภาพ และยังสามารถที่จะพัฒนาโปรแกรมได้ไม่ยากนัก หลักการของวิธีลิฟต์ได้รับการพัฒนามาจากการเรียนรู้โดยใช้กลุ่มตัวอย่าง (Instance-based Learning) ซึ่งอาศัยการถ่วงน้ำหนักของคุณลักษณะ [2] วิธีลิฟต์ใช้การหาคุณลักษณะที่มีความเกี่ยวข้องกับค่าเป้าหมาย (Target Concept) โดยอาศัยหลักการทางสถิติ ซึ่งเป็นวิธีการที่มีประสิทธิภาพ เวลาที่ใช้ในการทำงานของวิธีลิฟต์ขึ้นกับจำนวนของคุณลักษณะ และจำนวนตัวอย่างของข้อมูลที่ใช้ในการเรียนรู้ คุณลักษณะที่มีความเกี่ยวข้องกับค่าเป้าหมายจะถูกเลือก และจะละทิ้งคุณลักษณะที่ไม่มีความเกี่ยวข้องกับค่าเป้าหมายไป ทำให้

คุณลักษณะที่เหลืออยู่มีจำนวนลดลงได้ [3] ขั้นตอนการทำงานของวิธีลิฟต์แสดงในภาพที่ 1

Relief(S,m,t)

1. แยกตัวอย่าง S เป็นกลุ่มตัวอย่างบวก S+ และ กลุ่มตัวอย่างลบ S-
2. กำหนดเวกเตอร์น้ำหนักเริ่มต้น $W = (0, 0, \dots, 0)$
3. สำหรับ $i=1$ ถึง m
 - 3.1 เลือก $X \in S$ อย่างสุ่ม
 - 3.2 เลือกตัวอย่างบวกที่ใกล้กับ X มากที่สุด, $Z+ \in S+$
 - 3.3 เลือกตัวอย่างลบที่ใกล้กับ X มากที่สุด, $Z- \in S-$
 - 3.4 ถ้า X เป็นตัวอย่างบวก
แล้ว $\text{NearHit} = Z+$; $\text{NearMiss} = Z-$
มิฉะนั้น $\text{NearHit} = Z-$; $\text{NearMiss} = Z+$
 - 3.5 สำหรับ $j=1$ ถึง p เมื่อ p เป็นจำนวนคุณลักษณะ
คำนวณค่า W_j
 $W_j = W_j - \text{diff}(X_j - \text{NearHit}_j)^2 + \text{diff}(X_j - \text{NearMiss}_j)^2$
โดยที่ $\text{diff}(x,y) = \begin{cases} 0, & \text{if } x \text{ and } y \text{ are same} \\ 1, & \text{if } x \text{ and } y \text{ are different} \end{cases}$
4. $R = (1/m)W$
5. สำหรับ $j=1$ ถึง p
ถ้า $R_j \geq t$
แล้ว f_j เป็น คุณลักษณะที่เกี่ยวข้อง
มิฉะนั้น f_j เป็น คุณลักษณะที่ไม่เกี่ยวข้อง

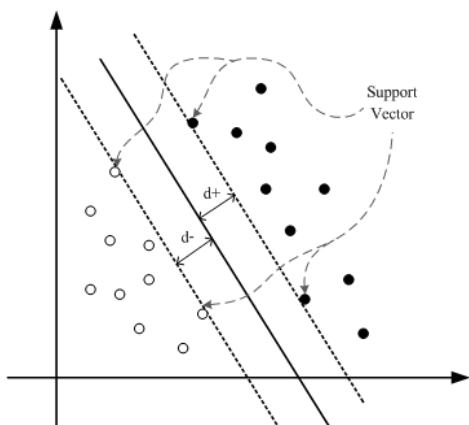
ภาพที่ 1: ขั้นตอนการทำงานของวิธีลิฟต์

วิธีลิฟต์จะทำการปรับค่าน้ำหนักของแต่ละคุณลักษณะ โดยใช้ระยะทางในการเลือกตัวอย่างใกล้ที่สุดที่อยู่ในกลุ่มเดียวกัน (Near-Hit) และ ตัวอย่างใกล้ที่สุดที่อยู่คนละกลุ่ม (Near-Miss) ผลลัพธ์ที่ได้จะเป็นค่าน้ำหนักที่อยู่ในช่วง -1 ถึง 1 ของแต่ละคุณลักษณะ [4] ในแต่ละรอบน้ำหนักของคุณลักษณะจะถูกปรับค่า และเมื่อครบตามจำนวนรอบที่กำหนดไว้หรือตรงตามเงื่อนไขที่ตั้งไว้ ก็จะพิจารณาค่าน้ำหนักของแต่ละคุณลักษณะว่าถึงค่าที่กำหนดหรือไม่ ถ้าถึงระดับที่กำหนดไว้ก็จะเลือกเก็บไว้เป็นคุณลักษณะที่มีความสำคัญต่อการจำแนกข้อมูลต่อไป จุดเด่นของวิธีลิฟต์ คือ ทนทานต่อสิ่งรบกวน แต่จะไม่ได้พิจารณาความสัมพันธ์ระหว่างกันของคุณลักษณะ

2.2 การจำแนกข้อมูลโดยใช้ซัพพอร์ตเวกเตอร์แมชชีน

ซัพพอร์ตเวกเตอร์แมชชีน หรือ SVM เป็นอัลกอริทึมที่เสนอโดย Vapnik [5], [6] ในปี 1982 เป็นเทคนิคการแบ่งกลุ่มข้อมูลโดยใช้หลักการของเส้นตรง (Linear Classifier) และเป็นเทคนิคหนึ่งที่ได้รับการนิยมน้อยมากในงานที่เกี่ยวข้องกับการจดจำรูปแบบ และการแก้ปัญหาการจัดกลุ่ม (Classification Problem) หลักการ SVM คือ การสร้างระนาบหลายมิติ (Hyperplane) ที่เหมาะสมบนข้อมูลตัวอย่าง (Training Data)

เพื่อแบ่งแยกกลุ่มข้อมูลที่แตกต่างกัน [7] ในการสร้างระนาบดังกล่าว ได้ทำการนิยามระยะห่างระหว่างจุดของข้อมูลที่อยู่ใกล้กับระนาบมากที่สุดทั้งสองด้าน คือ $d+$ และ $d-$ ระยะขอบ (Margin) เกิดจากระยะ $(d+) + (d-)$ ดังนั้นระนาบที่เหมาะสมคือระนาบที่มีค่าระยะขอบกว้างที่สุดดังแสดงในภาพที่ 2 โดยข้อมูลตัวอย่างที่อยู่บนขอบ จะถูกเรียกว่า ซัพพอร์ตเวกเตอร์ (Support Vector) [8]



ภาพที่ 2: การแบ่งกลุ่มข้อมูลโดยใช้ SVM

จากภาพเป็นการแบ่งกลุ่มข้อมูลออกเป็น 2 กลุ่ม โดยใช้ระนาบเชิงเส้น เพื่อให้อัลกอริทึมสามารถแบ่งแยกกลุ่มข้อมูลที่มีลักษณะไม่เป็นเชิงเส้น (Nonlinear Dataset) ได้ จำเป็นต้องแปลงกลุ่มข้อมูลตัวอย่างไปสู่ปริภูมิที่มีมิติสูงขึ้น ซึ่งถูกเรียกว่าปริภูมิแต่งเติม (Feature Space) โดยการแปลงดังกล่าวจะกระทำผ่านฟังก์ชันที่ไม่เป็นเชิงเส้น [8] และใช้การแก้ปัญหาลักษณะเพื่อหาสมการเชิงเส้นสำหรับแบ่งข้อมูลในปริภูมิแต่งเติมนั้น

ในงานวิจัยของ Dumais et. al. [9] ซึ่งทำการเปรียบเทียบกับวิธีการจำแนกเอกสาร โดยผลการทดสอบแสดงให้เห็นว่า SVM ให้ผลความถูกต้องในการจำแนกเอกสารที่สูงกว่าอัลกอริทึมอื่นเมื่อใช้เอกสารจำนวน 12,902 เอกสาร จากชุดข้อมูลทดสอบ Reuters-21578 และในงาน Joachims [10] ก็ได้เปรียบเทียบให้เห็นว่า SVM ให้ผลการทดสอบที่ดี เมื่อเปรียบเทียบกับเบย์อย่างง่าย โดยได้ทดสอบบนข้อมูล 2 ชุด คือ Reuters-21578 และ Ohsumed Corpus นอกจากนี้ยังมีงานวิจัยที่นำเอา SVM ไปประยุกต์ใช้เพื่อจำแนกข้อความเชิงความคิดเห็น [11] โดยใช้ SVM ในการพิจารณาว่าเป็นความคิดเห็นในเชิงบวกหรือเชิงลบ

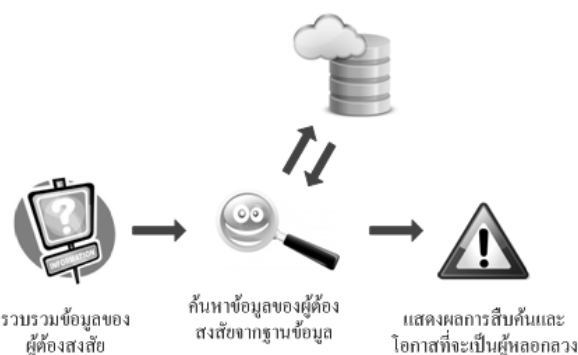
และมีการปรับเปลี่ยน Kernel Function ให้เหมาะสมกับการใช้งานด้วย

3. การกำหนดค่าน้ำหนักให้กับคุณลักษณะ

ดังได้กล่าวแล้วว่างานวิจัยนี้เสนอที่จะใช้วิธีเพื่อช่วยในการให้ค่าน้ำหนักสำหรับแต่ละคุณลักษณะในการสืบค้นข้อมูลของผู้ต้องสงสัยที่หลอกลวงขายสินค้าทางเครือข่ายอินเทอร์เน็ต ซึ่งมีลักษณะที่ข้อมูลอาจมาจากการปลอมแปลง ทำให้ค่าความน่าเชื่อถือของแต่ละคุณลักษณะมีค่าไม่แน่นอน ในหัวข้อนี้จะอธิบายถึงแนวคิดในการทำงานของระบบ คุณลักษณะที่จะนำมาใช้ในการสืบค้น และการให้ค่าน้ำหนักบนคุณลักษณะที่ใช้สืบค้น

3.1 การทำงานของระบบสืบค้น

ลักษณะการทำงานของระบบสืบค้น เมื่อผู้ใช้ระบบต้องการซื้อสินค้าหรือบริการผ่านทางอินเทอร์เน็ต โดยได้เลือกสินค้าหรือบริการที่ต้องการและมีข้อมูลของผู้ขายสินค้าแล้ว หากต้องการตรวจสอบว่าผู้ขายเคยมีประวัติหลอกลวงหรือไม่ หรือมีโอกาสสกนน้อยเพียงใดที่จะเป็นผู้หลอกลวงขายสินค้า ก็จะนำข้อมูลของผู้ขายสินค้ามาสืบค้นจากระบบ ภาพรวมการทำงานของระบบแสดงดังภาพที่ 3



ภาพที่ 3: การทำงานของระบบสืบค้น

ระบบจะทำการตรวจสอบข้อมูลกับฐานข้อมูลที่ได้รับความร่วมมือจากเว็บไซต์พันธมิตร ซึ่งจะแจ้งข้อมูลเมื่อเกิดกรณีร้องเรียนว่ามีกรหลอกลวงเกิดขึ้น ในการตรวจสอบจะมีการคำนวณค่าความคล้ายคลึงกันของข้อมูล โดยใช้ค่าน้ำหนักสำหรับแต่ละคุณลักษณะที่นำไปใช้ในการสืบค้นร่วมในการคำนวณความคล้ายคลึงด้วย ผลที่ได้ออกมาจะแสดงของผู้ที่เคยกระทำผิดที่มีข้อมูลคล้ายคลึงกับผู้ต้องสงสัย โดยเรียงลำดับความคล้ายคลึงจากมากไปน้อย

3.2 คุณลักษณะที่ใช้ในสืบค้นข้อมูลของผู้หลอกลวงขายสินค้าบนอินเทอร์เน็ต

จากการศึกษาข้อมูลการฉ้อโกง ทำให้ทราบว่าผู้ที่ซื้อสินค้า มักมีข้อมูลของผู้ขายเท่าที่จำเป็นต่อการติดต่อซื้อขายเท่านั้น อาทิเช่น ชื่อสมาชิกที่ใช้ในการเข้าสู่ระบบ (Username) หรือ ชื่อเล่น หมายเลขโทรศัพท์เคลื่อนที่ และ หมายเลขบัญชีธนาคาร ส่วนข้อมูลอื่นๆอาจมีการสอบถามบ้างในบางราย ผู้วิจัยจึงได้เลือกใช้ 10 คุณลักษณะของข้อมูลที่จะเป็นประโยชน์สำหรับการสืบค้นได้ และสามารถเก็บรวบรวมข้อมูลเหล่านั้นได้ ซึ่งได้แก่ (1) Username (2) ชื่อ (3) นามสกุล (4) หมายเลขบัตรประชาชน (5) เบอร์โทรศัพท์ (6) อีเมลล์ (7) ชื่อธนาคารที่โอนเงิน (8) สาขาของธนาคาร (9) หมายเลขบัญชีธนาคาร และ (10) หมายเลข IP

จากการสังเกต 10 คุณลักษณะดังกล่าว พบว่าคุณลักษณะส่วนใหญ่สามารถเปลี่ยนแปลงหรือปลอมแปลงได้ เช่น Username สามารถตั้งได้เอง หรือ สมัครสมาชิกใหม่เพื่อให้ได้ Username ใหม่ ชื่อและนามสกุลอาจไม่ใช่ข้อมูลจริง เบอร์โทรศัพท์และอีเมลล์ ถึงแม้จะเปลี่ยนแปลงได้ แต่ก็มีความถี่ที่มีการนำกลับมาใช้อีก หมายเลขบัญชีธนาคารมักจะมีการใช้ซ้ำอยู่เรื่อยๆ แต่ก็มีความถี่ที่เป็นการจ้างวานผู้อื่นเปิดบัญชี ส่วนหมายเลข IP มักเป็นข้อมูลที่ปลอมแปลงน้อยกว่าคุณลักษณะอื่นๆ แต่ก็ไม่ง่ายที่จะได้หมายเลข IP มา เนื่องจากบางเว็บไซต์ไม่มีนโยบายในการเปิดเผยหมายเลข IP ของสมาชิกในเว็บไซต์นั้น

3.3 การกำหนดค่าน้ำหนักบนคุณลักษณะที่ใช้ในการสืบค้น

ในการกำหนดค่าน้ำหนักให้กับแต่ละคุณลักษณะจะใช้วิธีการวิธีลิฟ เพื่อช่วยในการกำหนดค่าน้ำหนัก โดยจะนำตัวอย่างที่มีส่งให้วิธีลิฟเรียนรู้ค่าน้ำหนัก จากนั้นจะนำคุณลักษณะมาจัดเรียงตามค่าน้ำหนักที่ได้จากรีลิฟ จากมากไปน้อย เพื่อทำการแปลงค่าน้ำหนักเหล่านั้นให้เป็นเปอร์เซ็นต์ ค่าน้ำหนักจากรีลิฟที่น้อยที่สุด จะถูกแปลงเป็นค่าน้ำหนักที่ 1% เสมอ เพื่อให้การสืบค้นด้วยคุณลักษณะนั้นยังมีโอกาสพบข้อมูล ส่วนคุณลักษณะอื่นๆ สามารถคำนวณค่าน้ำหนักเป็นเปอร์เซ็นต์ได้จากสมการ (1)

$$P_i = \frac{(W_i - W_{min}) \times 99}{\sum_{i=1}^n (W_i - W_{min})}, \quad i = 1, 2, \dots, n - 1 \quad (1)$$

เมื่อ P_i เป็นค่าน้ำหนักที่เป็นเปอร์เซ็นต์,

W_i เป็นค่าน้ำหนักที่ได้จากรีลิฟ, และ

W_{min} เป็นค่าน้ำหนักที่น้อยที่สุดที่ได้จากรีลิฟ, และ n เป็นจำนวนคุณลักษณะทั้งหมดที่นำมาพิจารณา ซึ่ง $P_n = 1$

สำหรับข้อมูลที่ให้วิธีลิฟทำการเรียนรู้ได้ทำการเก็บรวบรวมข้อมูลของผู้ที่เคยมีประวัติกระทำการหลอกลวงจากเว็บไซต์ที่มีการประกาศขายสินค้าออนไลน์ ซึ่งเป็นเว็บไซต์ที่มีชื่อเสียงจำนวน 7 เว็บไซต์ โดยรวบรวมข้อมูลของผู้ขายสินค้าทั่วไปที่ไม่มีประวัติการฉ้อโกงจำนวน 100 ราย และ ผู้ที่มีประวัติฉ้อโกงจำนวน 100 ราย ผลการให้ค่าน้ำหนักแสดงดังตารางที่ 1

ตารางที่ 1: ผลการกำหนดค่าน้ำหนักให้กับแต่ละคุณลักษณะด้วยวิธีลิฟโดยใช้ตัวอย่างข้อมูลจริง

ลำดับ	คุณลักษณะ	ค่าน้ำหนักที่ได้จากรีลิฟ	ค่าน้ำหนักหลังจากปรับเป็นเปอร์เซ็นต์
1	Name	0.125 ± 0.003	19.5074
2	Surname	0.109 ± 0.004	17.9468
3	Telephone Number	0.035 ± 0.002	10.7291
4	Bank Number	0.027 ± 0.003	9.9488
5	Bank Branch	0.022 ± 0.002	9.4611
6	Username	0.011 ± 0.002	8.3882
7	E-mail	0.008 ± 0.002	8.0956
8	IP Address	0.003 ± 0.001	7.6079
9	National ID	0.000 ± 0.000	7.3153
10	Bank Name	-0.075 ± 0.002	1.0000
เปอร์เซ็นต์รวม			100.0000

จากผลการทดลองให้ค่าน้ำหนักพบว่า ชื่อ นามสกุล และ หมายเลขโทรศัพท์ เป็นคุณลักษณะที่มีความสำคัญเป็นอันดับต้นๆ ทั้งที่คุณลักษณะเหล่านี้สามารถปลอมแปลงได้ง่าย แต่คุณลักษณะสำคัญที่สามารถใช้ระบุตัวตนได้ทันทีเช่น IP Address หรือ National ID กลับอยู่อันดับต่างๆ เนื่องจากข้อมูลเหล่านี้ของผู้กระทำความผิดมักไม่ถูกเปิดเผย ทำให้มีค่าว่างในข้อมูลตัวอย่างเป็นจำนวนมาก ส่วนชื่อธนาคารมักเป็นข้อมูลที่มีเหมือนกันทั้งในผู้ฉ้อโกงและผู้ค้าปรกติจึงทำให้มีค่าน้ำหนักค่อนข้างต่ำ

เพื่อจะเลียนแบบพฤติกรรมของผู้กระทำความผิดที่มักจะมีการปลอมแปลงข้อมูล จึงได้ทำการสมมติข้อมูลโดยเลียนแบบพฤติกรรมการให้ข้อมูล ที่ผู้หลอกลวงมักจะกระทำ เช่น ปลอมแปลงชื่อ เปลี่ยนหมายเลขโทรศัพท์ หรือ เปลี่ยนหมายเลขบัญชีธนาคาร โดยได้ทำการจำลองข้อมูลขึ้นมาอีกรวมเป็น 408 รายการ แบ่งเป็นข้อมูลผู้ฉ้อโกงและไม่ใช่ผู้ฉ้อโกงอย่างละ 204 รายการ ผลการหาค่าน้ำหนักของข้อมูลชุดนี้แสดงในตารางที่ 2

เหตุที่ต้องจำลองข้อมูลขึ้นมาเพิ่มเติม เพราะจากข้อมูลที่เก็บรวบรวมมา ไม่สามารถระบุได้ว่าเป็นผู้ค้าคนเดิมที่มีการเปลี่ยนแปลงข้อมูลหรือไม่ การเลียนแบบพฤติกรรมของผู้ซื้อ/โกง ทำให้ได้ข้อมูลที่มากขึ้น และสามารถนำไปเรียนรู้ค่าถ่วงน้ำหนักที่เหมาะสมได้

ตารางที่ 2: ผลการกำหนดค่าน้ำหนักให้กับแต่ละคุณลักษณะด้วยวิธีฟิวส์โดยใช้ข้อมูลตัวอย่างที่สร้างขึ้น

ลำดับ	คุณลักษณะ	ค่าน้ำหนักที่ได้จากรีฟิวส์	ค่าน้ำหนักหลังจากปรับเป็นเปอร์เซ็นต์
1	National ID	0.045 ± 0.000	16.5825
1	IP Address	0.045 ± 0.000	16.5825
3	Bank Number	0.032 ± 0.002	13.3650
4	Bank Name	0.021 ± 0.004	10.6425
4	Surname	0.022 ± 0.001	10.8900
6	E-mail	0.015 ± 0.001	8.6625
6	Name	0.015 ± 0.001	9.1575
8	Telephone Number	0.008 ± 0.001	7.4250
9	Username	0.001 ± 0.000	5.6925
10	Bank Branch	-0.022 ± 0.004	1.0000
เปอร์เซ็นต์รวม			100.0000

ผลที่ได้จากการให้ค่าน้ำหนักด้วยวิธีฟิวส์เป็นไปตามคาดหมายคือ National ID และ IP Address เป็นคุณลักษณะสำคัญที่สามารถนำมาใช้ในการตรวจสอบข้อมูลของผู้ต้องสงสัยได้เป็นอย่างดี นั่นหมายความว่าหากเรารู้ National ID หรือ IP Address ของผู้ต้องสงสัย ในการสืบค้นข้อมูลก็ควรที่จะให้ค่าน้ำหนักกับคุณลักษณะเหล่านี้เพิ่มมากขึ้น

จุดประสงค์ของขั้นตอนนี้ไม่ได้ต้องการยืนยันว่า National ID หรือ IP Address เป็นคุณลักษณะที่ดี แต่ต้องการแสดงให้เห็นว่าวิธีฟิวส์สามารถสะท้อนลักษณะของข้อมูลที่ใช้สอนผ่านทางค่าน้ำหนักได้ เมื่อมีการเปลี่ยนแปลงชุดข้อมูลวิธีฟิวส์สามารถแสดงให้เห็นพฤติกรรมในข้อมูลที่เปลี่ยนแปลงไป ดังนั้นค่าน้ำหนักของวิธีฟิวส์จึงมีความเหมาะสมในการช่วยกำหนดค่าน้ำหนักของการสืบค้นข้อมูลได้

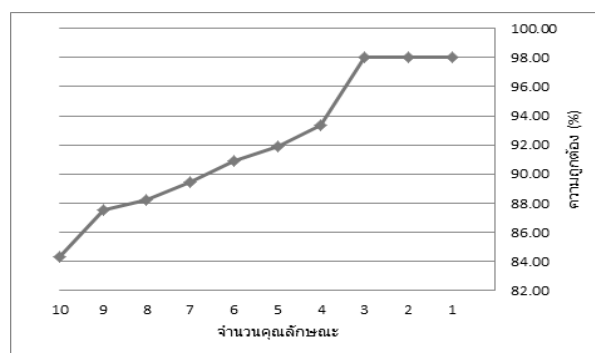
4. การทดสอบประสิทธิภาพของการกำหนดค่าน้ำหนัก

เพื่อที่ยืนยันประสิทธิภาพของวิธีฟิวส์ในการกำหนดค่าน้ำหนัก SVM ถูกนำมาใช้ในการทดสอบนี้ โดยจะทำการทดลองบนข้อมูลที่จำลองขึ้นจากการเลียนแบบพฤติกรรมที่เป็นไปได้ของผู้ซื้อ/โกง จำนวนคุณลักษณะจะถูกลดลงตามผลที่ได้จาก

วิธีฟิวส์ ซึ่งจะทำการตัดคุณลักษณะที่มีความสำคัญน้อยสุดออกไปก่อน และบันทึกผลที่ได้จากการจำแนกข้อมูล ผลการทดสอบแสดงดังตารางที่ 3 และ ภาพที่ 4

ตารางที่ 3: ความถูกต้องในการจำแนกข้อมูลด้วย SVM

จำนวนคุณลักษณะ	ค่าความถูกต้อง (%)	จำนวนคุณลักษณะ	ค่าความถูกต้อง (%)
10	84.3137	9	87.5000
8	88.2353	7	89.4608
6	90.9314	5	91.9118
4	93.3824	3	98.0392
2	98.0392	1	98.0392



ภาพที่ 4: กราฟแสดงผลการจำแนกด้วย SVM เมื่อลดจำนวนคุณลักษณะ

จากกราฟแสดงให้เห็นว่าค่าความถูกต้องเพิ่มมากขึ้นเมื่อทำการตัด สาขาของบัญชีธนาคาร ซึ่งมีระดับความสำคัญน้อยสุดออกไป และเมื่อตัดคุณลักษณะอื่นๆออก ค่าความถูกต้องเพิ่มขึ้นอย่างต่อเนื่อง ซึ่งสอดคล้องกับข้อสมมติ ที่คุณลักษณะเหล่านี้อาจไปซ้ำกับผู้ชายปกติที่ไม่ได้ซื้อ/โกง หรือเป็นข้อมูลที่ผู้ซื้อ/โกงสามารถเปลี่ยนแปลงได้ง่าย แต่เมื่อใช้หมายเลขบัญชีธนาคาร IP Address และ หมายเลขประจำตัวประชาชน ซึ่งเป็นข้อมูลที่ปลอมแปลงได้ยาก ผู้ซื้อ/โกงมีโอกาสยากกลับมาใช้ใหม่ได้มาก และมักไม่ซ้ำกับข้อมูลของผู้ชายที่ไม่ได้มีพฤติกรรมซื้อ/โกง การใช้ข้อมูลเหล่านี้จึงให้ความถูกต้องสูงกว่า

นอกจากนี้ยังได้ทดลองนำข้อมูลของผู้ที่เคยถูกแจ้ง หรือ ถูกสงสัยว่าเป็นผู้หลอกลวงจำนวน 187 ราย เก็บลงในฐานข้อมูล และใช้ข้อมูลผู้ต้องสงสัยอีกชุดหนึ่ง จำนวน 20 ราย นำมาทดลองสืบค้นจากฐานข้อมูล เปรียบเทียบผลการสืบค้นแบบใส่ค่าน้ำหนักที่ได้จากรีฟิวส์ กับผลการสืบค้นแบบกำหนดให้ทุกคุณลักษณะมีค่าน้ำหนักเท่ากันทั้งหมด ค่าน้ำหนักที่ได้จากการเปรียบเทียบกับผลการสืบค้นที่มีค่าน้ำหนักสูงสุดแสดงในตารางที่ 4

ตารางที่ 4: ค่าน้ำหนักที่ได้จากการสืบค้นแบบใส่ค่าน้ำหนักด้วยวิธีลิฟ

ตัวอย่างที่	น้ำหนักเท่ากัน	น้ำหนักวิธีลิฟ	ตัวอย่างที่	น้ำหนักเท่ากัน	น้ำหนักวิธีลิฟ
1	10.0000	24.6745	11	10.0000	18.3739
2	0	0	12	20.0000	19.9482
3	10.0000	26.2234	13	40.0000	45.0777
4	10.0000	24.6978	14	40.0000	45.0777
5	10.0000	22.6974	15	20.0000	35.3253
6	10.0000	16.1485	16	30.0000	42.0553
7	10.0000	15.2574	17	10.0000	23.9030
8	10.0000	23.0424	18	10.0000	23.9030
9	0	0	19	10.0000	16.8943
10	10.0000	18.3739	20	0	0

ค่าน้ำหนักที่เป็น 0 คือ ไม่พบข้อมูลที่สอดคล้องกับผู้ต้องสงสัยเลย เหตุที่ตัวเลขไม่สูงมากนักเนื่องจากข้อมูลที่เก็บอยู่ในฐานข้อมูลเป็นค่าว่างถึง 63.32% มีหลายกรณีที่มีแค่ ชื่อผู้ใช้งาน หมายเลขโทรศัพท์ และหมายเลขบัญชี ทำให้การสืบค้นข้อมูลทำได้ยาก การใช้ค่าน้ำหนักที่เท่ากันจะบอกได้เพียงแค่ว่าเหมือนกันก็คุณลักษณะ แต่ไม่สามารถบอกได้ว่าคุณลักษณะที่มีอยู่นั้นช่วยระบุตัวผู้ก่อโงได้มากน้อยเพียงใด เช่น ในตัวอย่างที่ 2-8 จะเห็นว่าหากใช้ค่าน้ำหนักเท่ากันจะได้ผลออกมาเท่ากันหมด แต่หากมีการใช้ค่าน้ำหนักจะพบตัวอย่างที่ 3 จะมีความน่าสงสัยว่าจะเป็นผู้ก่อโงมากกว่าตัวอย่างที่ 7 เป็นต้น

5. สรุป

บทความนี้ได้นำเสนอเทคนิคการสืบค้นข้อมูลของผู้กระทำการหลอกลวงขายสินค้าบนอินเทอร์เน็ต ซึ่งมักมีพฤติกรรมการปลอมแปลงข้อมูลบางอย่าง โดยได้ทำการกำหนดค่าน้ำหนักให้กับคุณลักษณะที่นำมาใช้ในการสืบค้น ค่าน้ำหนักที่ใช้ได้มากจากการเรียนรู้ด้วยวิธีลิฟ บนข้อมูลจริงและข้อมูลเพิ่มเติมที่สร้างขึ้นจากการจำลองพฤติกรรมการปลอมแปลงข้อมูลของผู้กระทำความผิด ซึ่งผลการทดสอบโดยใช้ซัพพอร์ตเวกเตอร์แมชชีนเป็นตัวจำแนกข้อมูลได้แสดงให้เห็นว่าค่าน้ำหนักที่ได้จากรีลิฟสอดคล้องกับพฤติกรรมของผู้ก่อโง จึงเป็นค่าน้ำหนักที่สามารถนำไปใช้ในกระบวนการสืบค้นได้

ข้อมูลสำคัญสำหรับการระบุตัวผู้ก่อโงจากการศึกษานี้ คือ หมายเลขบัตรประชาชน หมายเลขประจำเครื่อง และ หมายเลขบัญชีธนาคาร แต่หากขาดข้อมูลเหล่านี้ ข้อมูลอื่นๆที่ได้มาสามารถนำมาใช้ประกอบกันเพื่อใช้พิจารณาเปรียบเทียบในลักษณะที่ให้ความสำคัญกับแต่ละคุณลักษณะไม่เท่ากัน หรือ

ใส่ค่าถ่วงน้ำหนักได้ ซึ่งจะทำให้ความแม่นยำในการสืบค้นข้อมูลเพิ่มมากขึ้น จากการทดสอบการสืบค้นพบว่าได้ผลการสืบค้นที่มีประสิทธิภาพการดีกว่าการไม่ใช้ค่าน้ำหนัก ซึ่งสามารถนำไปประยุกต์ใช้กับข้อมูลที่ขาดความครบถ้วนได้เป็นอย่างดี

กิตติกรรมประกาศ

การศึกษานี้เป็นส่วนหนึ่งของโครงการ การวิเคราะห์ความสัมพันธ์และความเชื่อมโยงของข้อมูลเพื่อใช้ในการติดตามตัวผู้กระทำความผิด ซึ่งได้รับทุนอุดหนุนการวิจัยจากสำนักงานคณะกรรมการวิจัยแห่งชาติ ปีงบประมาณ 2556

เอกสารอ้างอิง

- [1] K. Kira and L. Rendell, "A practical approach to feature selection," *Proceeding of the ninth international workshop on machine learning*, Morgan Kaufmann Publisher Inc., pp. 249-256, 1992.
- [2] T. Kasemsrithanawat and T. Phienthrakul, "Sentiment Classification using Bayes' Classifier and Feature Selection with Relief Algorithm," *National Conference on Computer Information Technologies (CIT 2011 & UniNOMS 2011)*, NakornPathom, Thailand, 26-28 January 2011.
- [3] สุรศักดิ์ คังสกุล และ ธาปณี เสงสนั่นกุล, "การตัดคำภาษาไทยสำหรับข้อความในพีพธิภณซ์ปลาน้ำจืด," *วารสารวิทยาศาสตร์บูรพา*, ปีที่ 16, ฉบับที่ 1, หน้า 84-93, พ.ศ. 2554.
- [4] M. Robnik-Sikonja and I. Kononenko, "Theoretical and empirical analysis of ReliefF and RReliefF," *Machine Learning*, vol. 53, pp. 23-69, 2003.
- [5] V. Vapnik, "Estimation of dependences based on empirical data," Springer-Verlag, 1982.
- [6] V. Vapnik, "An Overview of Statistical Learning Theory," *IEEE Transactions on Neural Networks*, vol. 10, no. 5, pp. 988-999, 1999.
- [7] N. Cristianini and J. Shawe-Taylor, "An Introduction to Support Vector Machines and Other Kernel-based Learning Methods," Cambridge University Press, U.K., 2000.
- [8] C.J.C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition," *Data Mining and Knowledge Discovery*, vol. 2, no. 2, pp. 955-974, 1998.
- [9] S. Dumais, J. Platt, D. Heckman, and M. Sahami, "Inductive Learning Algorithms and Representations for Text Categorization," *CIKM '98 Proceedings of the seventh international conference on Information and knowledge management*, ACM New York, NY, USA, pp. 148-155, 1998.
- [10] T. Joachims, "Text Categorization with Support Vector Machines: Learning with Many Relevant Features," *Proceedings of the European Conference on Machine Learning (ECML)*, Springer, 1998.
- [11] T. Phienthrakul, B. Kijsirikul, H. Takamura, and M. Okumura, "Sentiment Classification with Support Vector Machines and Multiple Kernel Functions," *The 6th International Conference on Neural Information Processing (ICONIP 2009)*, Bangkok, Thailand, 1-5 December 2009.

การลดมิติข้อมูลการวิเคราะห์ความสัมพันธ์และการประยุกต์ สำหรับวิเคราะห์ข้อมูลพื้นฐานการใช้งานสมาร์ตโฟน

Data Dimension Reduction by Association and Application for Analysis Smart Phone Usage Data

พฤตพิงศ์ เพ็งศิริ(PrudtipongPongsiri)¹ สุริยะะ พินิจการ(Suriyapinitkan)² ณัฐชดา มงคลชาติ (NutchudaMongkolchart)³

นาวพร วิสิฐพงษ์พันธ์ (NawapornWisitpongphan)⁴ และ พยุง มีสัง (PhayungMeesad)⁵

¹คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ

²คณะศิลปศาสตร์และวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตกำแพงแสน

^{3,4,5} คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹prudtipong.p@hotmail.com, ²faassypi@ku.ac.th, ³lottehunterb18@gmail.com,

⁴nawapornn@kmutnb.ac.th, ⁵pym@kmutnb.ac.th

บทคัดย่อ

บทความนี้เป็นการเสนอวิธีใหม่ การวิเคราะห์ปัจจัยในการพัฒนาสมาร์ตโฟนด้วยการแบ่งส่วนข้อมูลเพื่อหาความสัมพันธ์แบบ Apriori เรียกว่า PAA (Partition data of Association with Apriori) และนำมาวัดประสิทธิภาพการลดจำนวนมิติข้อมูลกับวิธีคัดเลือกคุณลักษณะของข้อมูลในแบบต่างๆ เพื่อหาค่าความถูกต้อง ด้วยวิธีการจำแนกข้อมูลแบบโครงข่ายประสาทเทียม (ANN) ซึ่งผลการวิจัยสามารถสรุปได้ว่าการลดจำนวนมิติข้อมูลด้วยเทคนิคPAA นี้สามารถบ่งชี้ถึงกฎความสัมพันธ์ของข้อมูลที่เป็นปัจจัยด้านแนวทางการพัฒนาสมาร์ตโฟนอย่างมีนัยสำคัญ

คำสำคัญ: กฎความสัมพันธ์, สมาร์ตโฟน, อพริ ออริ, การคัดเลือกคุณลักษณะ, โครงข่ายประสาท

Abstract

This paper proposes a new method to analyze factors in the development of smart phones using partition data of association with Apriori (PAA). The PAA is compared with other feature selection methods in order to compare the effectiveness in reducing the dimension of the data. Accuracy is measured by means of adaptive neural network (ANN). Our results show that dimension which is reduced by PAA technique provide association rules of

factors that are significant in the development of smart phones.

Keyword: Smart phone, Association rule, Apriori, Feature selection, Neural network

1. บทนำ

เครือข่ายสังคม (Social network) [1] กลายเป็นปัจจัยหนึ่งที่สำคัญในการติดต่อสื่อสารในยุคโลกภิวัตน์ โดยข้อมูลที่อยู่ในเครือข่ายสังคมที่มีรูปแบบอันหลากหลาย และสามารถแบ่งปันข้อมูลเหล่านั้นให้กับสมาชิกบุคคลอื่นๆ ได้ จึงถือว่าการเพิ่มประสิทธิภาพในการสื่อสารได้เป็นอย่างดี โดยประชากรส่วนใหญ่ที่เป็นสมาชิกเข้าถึงเครือข่ายสังคมผ่านเครื่องคอมพิวเตอร์หรืออุปกรณ์สื่อสารเคลื่อนที่ ซึ่งในงานวิจัยนี้จะมุ่งเน้นอุปกรณ์สื่อสารเคลื่อนที่แบบสมาร์ตโฟน (Smart phone) เนื่องจากสามารถตอบสนองความต้องการเชื่อมต่อ ซึ่งเป็นการใช้งานทางเครือข่ายสังคมได้อย่างสะดวกและง่ายที่สุด การพัฒนาสมาร์ตโฟนจึงต้องคำนึงความสอดคล้องในการใช้งานเครือข่ายซึ่งเป็นส่วนที่สำคัญยิ่ง นอกจากนี้ส่วนที่สำคัญอีกประการหนึ่งคือ การปรับฟังก์ชันการใช้งานและแอปพลิเคชันของสมาร์ตโฟนจะต้องมีจุดดึงดูดความสนใจ และไม่ยากต่อการใช้งาน ดังนั้นจึงมีความจำเป็นต้องศึกษาปัจจัยในการใช้เครือข่ายสังคม

ของประชากร ซึ่งเป็นการส่งผลทำให้การพัฒนาสมาร์ตโฟนมีประสิทธิภาพและตอบสนองการใช้งานมากขึ้น

งานวิจัยนี้ได้มุ่งเน้นการวิเคราะห์ข้อมูลเครือข่ายสังคมของประชากรที่ใช้งานบนพื้นฐานสมาร์ตโฟน โดยการแบ่งกลุ่มการตอบสนองความต้องการในการใช้สมาร์ตโฟน และได้ทำการสำรวจจากแบบสอบถามเกี่ยวข้องกับการใช้เครือข่ายสังคม [2] และนำผลลัพธ์ที่ได้ขึ้นมาแบ่งระดับความสัมพันธ์ของแอททริบิวต์เพื่อเพิ่มความสามารถหรือฟังก์ชันของสมาร์ตโฟนที่บ่งชี้แนวทางการพัฒนาต่อไป

2. ทฤษฎีและวรรณกรรมที่เกี่ยวข้อง

การวิเคราะห์ถึงการพัฒนาสมาร์ตโฟนนั้นจำเป็นต้องทราบถึงความต้องการใช้งานเครือข่ายสังคม ซึ่งมีการเชื่อมต่อแบบที่มีมีการเคลื่อนไหวอยู่ตลอดเวลา ดังนั้นผู้วิจัยจึงได้นำเสนอกรอบความคิดและทำการศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้องดังนี้

2.1 การวิเคราะห์แบบสอบถามเพื่อแบ่งส่วนข้อมูล

เป็นการวิเคราะห์จากข้อมูลแบบสอบถาม [3] ที่เกี่ยวข้องกับเครือข่ายสังคม ซึ่งผู้วิจัยได้แบ่งเป็น 2 ส่วนคือ ส่วนแรกการใช้งาน (Usage) ส่วนที่สองคือสมาร์ตโฟนและระบบปฏิบัติการ (Smart phone and Operating system) และแอปพลิเคชัน ในการวิเคราะห์นี้ผู้วิจัยได้ดำเนินการแบ่งส่วนข้อมูล (Partition of data) และการตัดลดแอททริบิวต์ (Pruning attribute) [4], [5] ที่ไม่มีความสัมพันธ์หรือเป็นเขตว่างออกไป

2.2 กฎความสัมพันธ์ (Association rule) [6], [7]

เป็นหนึ่งในเทคนิคของ Data mining ที่บ่งบอกถึงความสัมพันธ์ในการเปรียบเทียบข้อมูล ซึ่งจะต้องมีการตั้งกฎเพื่อนำไปวิเคราะห์ในการตัดสินใจ เช่นเดียวกันหนึ่งในอัลกอริทึมของกฎความสัมพันธ์นั้นคือ อัลกอริทึมอปริออริ (Apriori algorithm) [4] ซึ่งจะทำการคำนวณหาความสัมพันธ์ของ Itemsets โดยความสัมพันธ์ ของ Itemsets นั้นเรียกว่า กฎความสัมพันธ์ (Association rule) ซึ่งจะอยู่ในรูปแบบดังต่อไปนี้

- $\{Item1, Item2\} \rightarrow \{Item3\}$ คือกฎความสัมพันธ์ที่ถูกสร้างขึ้นมา
- ค่าสนับสนุน คือ $Support = P(x \cap y)$
- ค่าความเชื่อมั่น คือ $Confidence = \frac{P(xy)}{P(y)}$
- เมื่อ x และ y เป็นสมาชิกของ Itemset

Itemset คือเซตของรายการที่มีความยาวหรือจำนวนรายการในอัลกอริทึมอปริออริ จะต้องมีการกำหนดค่าสนับสนุนขั้นต่ำ (Minimum support หรือ Minsup) และค่าความเชื่อมั่นขั้นต่ำ

(Minimum confidence หรือ Minconf) ซึ่งในการกำหนดค่าทั้งสองนี้จะขึ้นอยู่กับผู้ใช้ระบบเป็นผู้กำหนดเอง

2.3 CBA อัลกอริทึม

Classification Based on Associations (CBA) [7] เป็นการเสนอวิธีการรวมอัลกอริทึมในการสืบค้นกฎความสัมพันธ์ เข้ากับเทคนิคการจำแนกประเภทข้อมูล โดยเป็นอัลกอริทึมแบบ CBA ประกอบด้วย 2 ส่วนคือ ส่วนที่หนึ่งการสร้างกฎความสัมพันธ์ (Rule generator) โดยพิจารณาเฉพาะกฎความสัมพันธ์จำแนกประเภทข้อมูลที่มีความถี่ผ่านค่า Minsup และค่า Minconf และส่วนที่สองการสร้างโมเดลในการทำนายข้อมูล (Classifier builder)

2.4 การคัดเลือกมิติข้อมูล (Feature selection)

Feature selection [8] การเลือกแอททริบิวต์ที่มีความสำคัญน้อยออก เพื่อคุณภาพในการทำนายหลังจากที่ได้ตัดแอททริบิวต์บางตัวออก ซึ่งส่วนใหญ่จะให้ค่าความถูกต้องสูงขึ้น เพราะแอททริบิวต์ที่เหลือเป็นแอททริบิวต์โดยมีความสำคัญในงานวิจัยได้นำเสนอการลดมิติข้อมูลด้วยความสัมพันธ์ด้วยวิธีอัลกอริทึมอปริออริ ซึ่งได้นำมาเปรียบเทียบกับวิธีการคัดเลือกมิติข้อมูล โดยเลือกวิธีดังต่อไปนี้

2.4.1 การเลือกคุณสมบัติบนพื้นฐานความสัมพันธ์

การเลือกคุณสมบัติของแอททริบิวต์นั้น ต้องใช้การพิจารณาบนพื้นฐานความสัมพันธ์ (Correlation-based feature selection : Cfs) [9] เป็นการหากลุ่มคุณลักษณะที่ได้รับการประเมินค่าจากความสามารถในการคาดการณ์ โดยคุณลักษณะที่ถูกคัดเลือกใช้สำหรับจำแนกประเภทของข้อมูล สามารถจัดการกับ คุณลักษณะที่ไม่เกี่ยวข้องกันที่จะให้ประสิทธิภาพในการจำแนกประเภทข้อมูลที่ต่ำคำนวณได้จากสมการ (1)

$$Merit_s = \frac{k\overline{r}_{cf}}{\sqrt{k+k(k-1)\overline{r}_{ff}}} \quad (1)$$

$Merit_s$ คือ Heuristic ที่เป็นกลุ่มของ Attribute S ที่ระกอดด้วย K Attribute ที่ถูกคัดเลือก

\overline{r}_{cf} คือ ค่าเฉลี่ยของกลุ่มของ Attribute ที่ถูกคัดเลือกของความสัมพันธ์ที่เกี่ยวข้องกับประเภทของข้อมูล

\overline{r}_{ff} คือ ค่าเฉลี่ยของกลุ่ม Attribute ที่ถูกคัดเลือกจากความสัมพันธ์ที่เกี่ยวข้องกันภายในกลุ่ม Attribute ที่ถูกคัดเลือก

2.4.2 Information Gain [10]

การพิจารณาจากค่าความน่าจะเป็นของแต่ละลักษณะที่เป็นไปได้แล้ววัดค่าความไร้ระเบียบ (Entropy) เพื่อคัดเลือกถ้าลักษณะใดให้ค่าเกณฑ์ที่สูงที่สุด แสดงว่าลักษณะนั้นสามารถ

จำแนกกลุ่มได้ดีที่สุดซึ่ง Information Gain มีการคำนวณที่ไม่ซับซ้อนแต่ยังมีความเอนเอียงในการประเมิน

Gain Ratio Criterion คำนวณอัตราส่วนเกณฑ์เป็นวิธีประเมินลักษณะที่ใช้วิธีเช่นเดียวกับ Information Gain แต่เพิ่มการหาค่าสารสนเทศการแบ่งแยก (Split Information) เพื่อแก้ไขความเอนเอียงซึ่งจะมีความละเอียดมากขึ้นโดยวิธีการคำนวณค่ามาตรฐานอัตราส่วนเกณฑ์สามารถแสดงได้ดังสมการ (2)

$$E(s) = -\sum_{c=1}^N p(S_c) \times \log_2 p(S_c) \quad (2)$$

วัดค่า Information Gain เพื่อสร้างลำดับ ดังสมการ(3)

$$Gain(S, V) = E(S) - \sum_{v \in \text{value}(v)} \frac{|S_v|}{S} \times E(S_v) \quad (3)$$

ลดค่าความเอนเอียง ดังสมการ (4)

$$SplitInfo(S, V) = \sum_{i=1}^m -\frac{|S_i|}{S} \times \log_2 \frac{|S_i|}{S} \quad (4)$$

หาค่า Gain Ratio Criterion ด้วยค่าการ Split Info ดังสมการ (5)

$$GainRatio(S, V) = \frac{Gain(S, V)}{SplitInfo(S, V)} \quad (5)$$

2.4.3 Relief algorithm

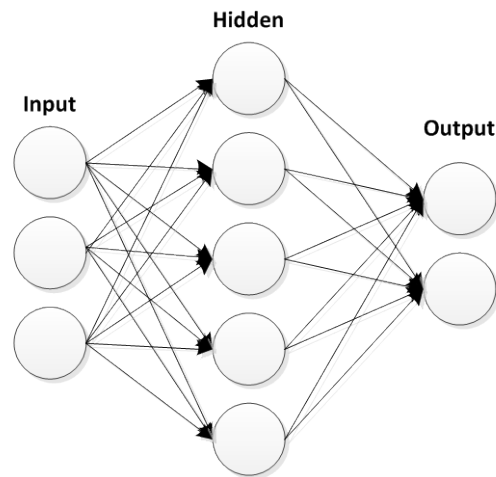
Relief [9],[11] เป็นอัลกอริทึมที่ได้รับความสนใจต่อเนื่องมาจากการเรียนรู้ของ Instance-based Learning ซึ่งได้ใช้การถ่วงน้ำหนักของคุณลักษณะ Relief โดยอาศัยหลักการทางสถิติและมีประสิทธิภาพของการทำงานที่ดี การทำงานของ Relief ใช้เวลาเป็นเชิงเส้นตรง (Linear Time) ซึ่งขึ้นกับจำนวนของคุณลักษณะ ดังสมการ (6)

$$Relief_x = \frac{P(\text{different value of } X | \text{different class})}{-P(\text{different value of } X | \text{same class})} \quad (6)$$

2.5 การจำแนกข้อมูล (Classification data)

การจำแนกประเภทข้อมูลนี้ เป็นการวิเคราะห์เพื่อจำแนกข้อมูลเป็นประเภทของกลุ่มข้อมูล โดยเทคนิคที่ใช้กับจะเป็นรูปแบบข้อมูลที่มีผู้สอน (Supervised Learning) โดยได้ศึกษา งานวิจัยที่เกี่ยวข้องคือ โครงข่ายประสาทเทียมแบบหลายชั้น (Multilayer Perceptron Artificial Neural Network: MLP) [12], [13] โดยลักษณะของโมเดลจะเป็นแบบหลายชั้น มีทั้งหมด 3 Layer ได้แก่ 1) Input Layer คือชั้นของการรับข้อมูลเข้าเพื่อทำการเรียนรู้ ในชั้นนี้จะมีการกำหนดจำนวนข้อมูลนำเข้า (Input) 2) Hidden Layer เป็นชั้นของการเรียนรู้ข้อมูลโดย Hidden

Layer ชั้นที่ 1 จะสุ่มค่าน้ำหนัก (Weight) ของข้อมูลนั้นๆ เพื่อเป็นการให้ความสำคัญของแต่ละข้อมูลที่ไม่เท่ากัน และสุ่มค่าความโน้มเอียง (Bias) เพื่อเป็นตัวกำหนดทิศทางการเรียนรู้ของโมเดล 3) Output Layer เป็นชั้นของข้อมูลผลลัพธ์จากการเรียนรู้ของโมเดล โดยข้อมูลนำเข้าของชั้นนี้คือข้อมูลผลลัพธ์จาก Hidden Layer ซึ่งจะมีการปรับค่าความโน้มเอียงและค่าน้ำหนักสามารถแสดงได้ดังภาพที่ 1



ภาพที่ 1: โครงข่ายประสาทเทียมแบบหลายชั้น

3. วิธีการดำเนินงานวิจัย

ในงานวิจัยนี้ได้นำเสนอกระบวนการวิเคราะห์ปัจจัยในการพัฒนาสมรรถนะโทรคมนาคมจากความสัมพันธ์ด้วยอัลกอริทึมอพริโอริ ซึ่งเป็นรายละเอียดของความถี่และความถี่ของสัญญาณ Mobile ในการเข้าถึงบริการเครือข่ายสังคม โดยมีกระบวนการในการวิเคราะห์ความสัมพันธ์และเปรียบเทียบประสิทธิภาพในการลดมิติข้อมูลดังนี้

3.1 การเตรียมข้อมูล (Pre-Processing)

จากแบบสอบถามในแหล่งข้อมูลเบื้องต้น สามารถทำการวิเคราะห์ข้อมูลเหล่านั้นเพื่อเตรียมการประมวลผลได้โดยผู้วิจัยได้มุ่งเน้นไปในส่วนของข้อมูลปัจจัยในการใช้แอปพลิเคชัน ซึ่งมีขั้นตอนการเตรียมข้อมูลดังนี้ 1) การกำจัด Missing data ทำให้ได้ข้อมูลที่มีส่วนสำคัญเท่านั้น 2) แยกข้อมูลจากแบบสอบถามที่อยู่ในหัวข้อเดียวกัน หรือมีความใกล้เคียงออกเป็นกลุ่มย่อย (Partition data) 3) ทำการแปลงข้อมูล (Data transform) โดยที่ใส่ค่าของค่าตอบกับ "1" ซึ่งเป็นเหตุผลหลัก และตัดค่าตอบอื่นทิ้งแบบ "Don't care" เนื่องจากค่าตอบอื่นมีลักษณะบ่งชี้ไปในทางการปฏิเสธ ซึ่งไม่เป็นประโยชน์ในการหาความสัมพันธ์

ของข้อมูลจึงทำให้มีข้อมูลเป็น 2 กลุ่มหลักมีทั้งหมด 688 เรคคอร์ด ซึ่งแบ่งเป็นกลุ่มที่เป็นฟังก์ชันหลักในการใช้งาน Social network มีจำนวน 17 แอททริบิวต์ และกลุ่มที่เป็นความต้องการและพฤติกรรมของผู้ใช้ Mobile มีจำนวน 7 แอททริบิวต์

โปรแกรมที่ใช้ในงานวิจัยครั้งนี้ ผู้วิจัยเลือกใช้โปรแกรม WEKA เวอร์ชัน 3.6.10 และ MATLAB ซึ่งเป็นซอฟต์แวร์ด้านการทำเหมืองข้อมูล และได้รับการยอมรับอย่างแพร่หลายทำการคัดเลือกข้อมูล (Feature selection) โดยอัลกอริทึมในการเลือกข้อมูลของโปรแกรม WEKA แบบ CfsSubsetEval, InfoGainAttributeEval และ ReliefAttributeEval และใช้โปรแกรม MATLAB ในการเปรียบเทียบประสิทธิภาพการลดมิติข้อมูลเพื่อหาค่าความถูกต้องจากกระบวนการเหล่านั้น ด้วยเทคนิคการจำแนกข้อมูลแบบ ANN (Adaptive Neural Network)

3.2 การสร้างโมเดลสำหรับการวิเคราะห์ความสัมพันธ์ของข้อมูลและการวัดประสิทธิภาพการลดมิติข้อมูล

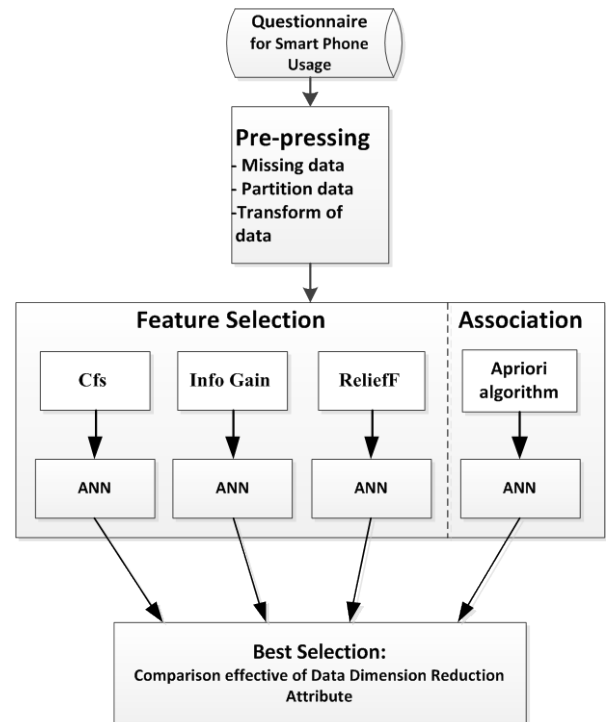
หลังจากผ่านกระบวนการเตรียมข้อมูลก่อนประมวลผลจากแหล่งข้อมูลแบบสอบถาม ผู้วิจัยได้สร้างโมเดลกระบวนการสำหรับการวิเคราะห์ข้อมูล และการวัดประสิทธิภาพการลดมิติข้อมูลโดยนำเสนอไว้ 3 ส่วนดังนี้

3.2.1 กระบวนการวิเคราะห์ปัจจัยในการพัฒนาสมาร์ตโฟน ด้วยกฎความสัมพันธ์ของข้อมูลแบบอปริโอริ เป็นการนำข้อมูลที่ได้อัตโนมัติจากการ Pre-process ทั้งสามขั้นตอนที่ได้แบ่งส่วนข้อมูล จึงเรียกวิธีนี้ว่า PAA หรือ Partition data of Association with Apriori เพื่อเข้าสู่กระบวนการหาความสัมพันธ์ (Association Rules) แบบอปริโอริ โดยได้ทำการแบ่งข้อมูลเป็นกลุ่มย่อย ซึ่งมีกลุ่มฟังก์ชันหลักในการใช้งาน Social network และกลุ่มความต้องการและพฤติกรรมของผู้ใช้ Mobile ซึ่งผู้วิจัยได้ตั้งสมมติฐานว่า ผลลัพธ์ที่ได้มาจากการกระบวนการนี้จะทำให้สามารถบ่งชี้ให้เห็นถึงการลดจำนวนมิติข้อมูลได้

3.2.2 กระบวนการลดมิติข้อมูลด้วยวิธีเลือกคุณสมบัติ (Feature Selection) แบบ Correlation-based Feature Selection, Information Gain และ Relief ข้อมูลที่นำเข้าสู่กระบวนการนี้เป็นข้อมูลเช่นเดียวกับในข้อ 3.2.1 แต่กระบวนการนี้เป็นการหาผลลัพธ์ ซึ่งได้มาจากการคัดเลือกคุณลักษณะในแต่ละแอททริบิวต์หรือมิติข้อมูล และจะทำให้ได้แอททริบิวต์ที่แตกต่างกันออกไปตามวิธีในแต่ละขั้นตอน

3.2.3 กระบวนการวัดและเทียบประสิทธิภาพระหว่างกฎความสัมพันธ์แบบ Apriori และการเลือกคุณสมบัติแบบ

ต่างๆ ซึ่งในกระบวนการนี้เป็นการวัดประสิทธิภาพของการลดมิติข้อมูลจากผลลัพธ์ในข้อ 3.2.1 และ 3.2.2 และเป็นการทดสอบเพื่อพิสูจน์ตามสมมติฐานที่ตั้งขึ้นมา ซึ่งโมเดลทั้ง 3 ส่วนนี้มีการทำงานที่เชื่อมโยงกัน ดังแสดงไว้ในภาพที่ 2



ภาพที่ 2 :โมเดลของกฎความสัมพันธ์และการวัดประสิทธิภาพการลดมิติข้อมูล

4. ผลการดำเนินงาน

การประเมินผลการทดสอบตามโมเดลที่ผู้วิจัยได้นำเสนอทั้ง 3 ส่วน สามารถประเมินและอภิปรายผลการทดสอบได้ดังนี้

4.1 ผลการทดสอบความสัมพันธ์ของข้อมูลด้วยอปริโอริ

การทดสอบในส่วนนี้ เป็นการหาความสัมพันธ์ของข้อมูลด้วยอปริโอริ ซึ่งจะเป็นการบ่งชี้ให้เห็นถึงทิศทางการพัฒนาสมาร์ตโฟนจากข้อมูลแบบสอบถาม โดยมีข้อมูลที่เน้นด้านการพัฒนาแอปพลิเคชันจำนวน 24 แอททริบิวต์และแบ่งขั้นตอนการทดสอบเป็น 2 ขั้นตอนดังนี้

1) ขั้นตอนในการหาความสัมพันธ์ของฟังก์ชันหลักในการใช้งาน Social network มีจำนวน 17 แอททริบิวต์ (แทนด้วยตัวแปร A,B,C,...,Q) ได้กำหนดค่าพารามิเตอร์ไว้ดังตารางที่ 1 ซึ่งผลการทดสอบพบว่า ได้กฎความสัมพันธ์ของข้อมูลจำนวน 14 กฎ และได้แอททริบิวต์ที่เกี่ยวข้องกันในกฎความสัมพันธ์นี้จำนวน 9 แอททริบิวต์โดยมีความหมายของความสัมพันธ์คือ

ผู้ใช้งาน Social network เป็นการใช้งานผ่าน MySpace, Facebook หรือ LnkedIn เพื่อหาข่าว สามารถจำกัดสิทธิ์ความเป็นเพื่อนที่เข้าถึง ซึ่งสามารถเปิดให้ทำการติดต่อกับสมาชิกครอบครัวหรือเพื่อนที่ไว้ใจได้ ต้องการให้ความสามารถของ Social network และทำให้มีการลบเพื่อนในลิสต์และย้ายรูปที่ถูกลบออกแต่ก็ทิ้งไปได้

ตารางที่ 1 : กำหนดข้อมูลที่ถูกรวบรวมในส่วนที่ 1

No	Attribute	Detail
1	A	ได้รับข่าวออนไลน์
2	B	ใช้ Social network ในรูปแบบ MySpace, Facebook or LnkedIn
3	C	รูปแบบ Profile ที่หลากหลาย
4	D	ความเป็นสาธารณะหรือส่วนตัว
5	E	จำกัดในสิ่งที่เพื่อน ๆ สามารถมองเห็นบางอย่าง
6	F	ใช้งานยากมาก
7	G	ลบคนที่มาจากเครือข่ายหรือรายชื่อเพื่อนของคุณ
8	H	Remove ชื่อของคุณจากภาพถ่ายที่ได้รับการติดแท็กเพื่อระบุดูคุณ
9	I	ลบความคิดเห็นที่คนอื่น ๆ ได้ทำในโปรไฟล์ของคุณ
10	J	การปรับปรุงการโพสต์ความเห็นรูปถ่ายหรือวิดีโอที่คุณเสียดใจในภายหลังการแบ่งปัน
11	K	ตั้งคำบัญชีของคุณเพื่อให้มันจะรวมถึงตำแหน่งของคุณบนข้อความของคุณ
12	L	สามารถติดต่อกับสมาชิกในครอบครัว
13	M	สามารถติดต่อกับเพื่อน ๆ ในปัจจุบัน
14	N	เชื่อมต่อกับเพื่อนเก่าที่คุณสูญเสียการติดต่อให้ใช้งานได้
15	O	สร้างเพื่อนใหม่
16	P	การเชื่อมต่อกับคนอื่น ๆ ที่มีงานอดิเรกหรือความสนใจของคุณ
17	Q	อ่านความคิดเห็นหรือโพสต์โดยคนคั้งการเมืองหรือนักกีฬา

ตารางที่ 2 : กำหนดข้อมูลที่ถูกรวบรวมในส่วนที่ 2

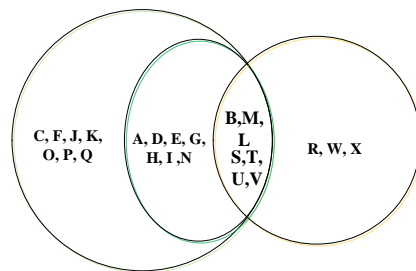
No	Attribute	Detail
1	R	กำลังมองหาสุขภาพหรือข้อมูลทางการแพทย์ออนไลน์
2	S	ดูข่าวออนไลน์สำหรับข้อมูลเกี่ยวกับการเมือง
3	T	ซื้อสินค้าออนไลน์เช่นหนังสือเพลงของเล่นหรือเสื้อผ้า
4	U	ซื้อหรือสำรองห้องพักออนไลน์สำหรับการบริการด้านการท่องเที่ยว เช่นเดียวกับตัวสายการบินห้องพักโรงแรมหรือรถเช่า
5	V	ดูออนไลน์สำหรับข้อมูลเกี่ยวกับงาน
6	W	ทำธุรกรรมธนาคารออนไลน์
7	X	เยี่ยมชมท้องถิ่นรัฐหรือรัฐบาลกลางเว็บไซต์ของรัฐบาล

2) เพิ่มข้อมูลในกลุ่มความต้องการและพฤติกรรมของผู้ใช้ Mobile จำนวน 7 แอททริบิวต์ เข้ารวมกับกลุ่มแรกเป็น 24 แอททริบิวต์ และกำหนดค่า minSup เพิ่มขึ้นเป็น 0.4 ซึ่งแสดงข้อมูลดังตารางที่ 2 เพราะจะแสดงถึงความสัมพันธ์ในแต่ละแอททริบิวต์ที่เกี่ยวข้องกันให้ได้มากที่สุด ผลการทดสอบพบว่าได้กฎความสัมพันธ์ของข้อมูลเพิ่มขึ้นเป็นจำนวน 30 กฎ และได้แอททริบิวต์ที่เกี่ยวข้องกัน ในกฎความสัมพันธ์นี้ลดลงเป็นจำนวน 9 แอททริบิวต์ ซึ่งมีความหมายของความสัมพันธ์ของกลุ่มนี้คือ ผู้ใช้งาน Social network เป็นการใช้งานผ่าน MySpace, Facebook หรือ LnkedIn ซึ่งเปิดให้ติดต่อกับครอบครัวและ

เพื่อนที่ไว้ใจได้ นอกจากนี้ยังใช้ Social network ในการติดตามข่าวสารการเมือง การซื้อของ เช่น หนังสือ เพลง ของเล่นต่างๆ รวมไปถึงการเรียกใช้บริการการท่องเที่ยว ซึ่งเชื่อมโยงกับธุรกิจของสายการบิน การบริการห้องพัก และสามารถติดตามข่าวสารการสมัครงาน และการทำงานได้ด้วยผลลัพธ์ของความสัมพันธ์ทั้ง 2 ชั้นตอนนี้ได้นำเสนอความสัมพันธ์ดังตารางที่ 3 และแสดงเป็นแผนภาพเซตข้อมูลแสดงดังภาพที่ 3

ตารางที่ 3 : ความสัมพันธ์ของข้อมูลด้วยฟรื่อรี

Issues	No.	Attr.	min	minSup	No.	No.	Asso.
	Attr		Con		rule	Attr	Attr
ฟังก์ชันหลักในการใช้งาน Social network	17	A,B,C,D,E,F,G,H,I,J,K,L,M,N,O,P,Q	0.9	0.1	14	9	A,B,D,E,G,H,L,M,N
ความต้องการและพฤติกรรมของผู้ใช้ Mobile	24	A,B,C,D,E,F,G,H,I,J,K,L,M,N,O,P,Q,R,S,T,U,V,W,X	0.9	0.4	30	7	B,L,M,S,T,U,V



ภาพที่ 3 : แผนภาพความสัมพันธ์ของข้อมูลแนวทางการพัฒนาสมาร์โฟน

4.2 ผลการทดสอบการวัดประสิทธิภาพการลดมิติข้อมูล

ผู้วิจัยจึงได้นำเสนอการเปรียบเทียบประสิทธิภาพการลดมิติข้อมูลตามโมเดลในข้อ 3.2.2 และ 3.2.3 โดยต้องผ่านกระบวนการจำแนกข้อมูลด้วย ANN ซึ่งได้เลือกแอททริบิวต์ข้อมูลมาทำการทดสอบเป็นจำนวน 17 แอททริบิวต์ ทำให้ได้ผลลัพธ์ที่ได้จะแสดงไว้ดังตารางที่ 2

ตารางที่ 4 : การวัดประสิทธิภาพการลดมิติข้อมูล

Method.	Attributes																	No. Attr.	MSE	Acc. (%)
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q			
PAA	✓	✓	-	✓	✓	-	✓	✓	✓	-	-	✓	✓	✓	-	-	-	9	0.159	84.1
Cfs	-	-	✓	✓	✓	✓	-	-	✓	✓	✓	✓	✓	✓	-	✓	✓	12	0.1552	84.48
InfoGain	-	-	-	✓	✓	✓	-	-	✓	-	-	✓	✓	✓	-	✓	✓	9	0.1954	80.46
ReliefF	-	-	-	-	✓	✓	✓	✓	-	✓	✓	✓	✓	✓	✓	✓	✓	9	0.1743	82.57
Count	1	1	1	3	3	3	2	2	4	1	2	4	4	4	1	3	3	-	-	-

ผลการทดสอบในตารางที่ 4 พบว่าประสิทธิภาพในการลดจำนวนมิติข้อมูลจากความสัมพันธ์ของข้อมูลด้วย PAA นั้น

สามารถลดจำนวนมิติข้อมูลจากการสุ่มเลือก 17 แอททริบิวต์ให้เหลือเป็นจำนวน 9 แอททริบิวต์ โดยมีค่าความถูกต้อง (Accuracy) อยู่ที่ 84.1% ซึ่งมีค่าความถูกต้องใกล้เคียงกับการเลือกใช้วิธีการเลือกคุณสมบัติแบบ Cfs เป็น 84.48% แต่ Cfs ลดจำนวนมิติข้อมูลได้เป็นจำนวน 12 แอททริบิวต์ ซึ่งมากกว่าแบบ PAA และได้เลือกมิติข้อมูลที่ตรงกันเป็นจำนวน 6 แอททริบิวต์ เช่นเดียวกับวิธีการเปรียบเทียบ Apriori กับ InfoGain และ ReliefF จะพบว่ามีความถูกต้องที่มากกว่า โดยที่มีการลดมิติข้อมูลที่เท่ากันคือ 9 แอททริบิวต์ ประสิทธิภาพของการลดจำนวนมิติข้อมูลแบบ PAA นั้นจะเห็นได้ว่ามีการเลือกมิติข้อมูลที่ตรงกันเมื่อเทียบกับการเลือกคุณสมบัติทั้ง 3 แบบเป็นจำนวน 4 แอททริบิวต์ คือ I, L, M และ N ซึ่งสามารถบ่งชี้ได้ว่าในมิติข้อมูลที่เลือกมามีความสัมพันธ์ โดยสอดคล้องกันเกี่ยวข้องกัน ซึ่งสามารถบ่งชี้ให้เห็นถึงแนวโน้มของเซตข้อมูลที่นำมาทดสอบ

5. สรุปและข้อเสนอแนะ

งานวิจัยนี้เป็นการวิเคราะห์ได้ว่าปัจจัยที่สำคัญต่อการพัฒนาสมาร์ตโฟนจากข้อมูลแบบสอบถามพื้นฐาน โดยได้นำเสนอวิธีการหาความสัมพันธ์ของข้อมูลด้วยฟอร์อิริ เพื่อเป็นการบ่งชี้ให้ทราบถึงปัจจัยในการพัฒนาโดยอาศัยการใช้งาน และรับรู้ผ่าน Social network ของผู้ใช้ การตอบสนองและการเข้าถึงข้อมูลที่เอื้อประโยชน์แก่การใช้ชีวิตประจำวัน ผู้วิจัยจึงได้นำเสนอวิธี PAA มาเป็นทางเลือกในการวิเคราะห์ความสัมพันธ์กับข้อมูลเหล่านี้ อีกทั้งได้ทำการวัดประสิทธิภาพการลดมิติข้อมูลเมื่อเปรียบเทียบกับวิธีการเลือกคุณสมบัติในรูปแบบ Cfs, InfoGain และ ReliefF โดยผ่านการจำแนกข้อมูลด้วย ANN ซึ่งพบว่าวิธีแบบ PAA นั้นมีค่าความถูกต้องใกล้เคียงกับ Cfs แต่ได้จำนวนมิติข้อมูลที่น้อยกว่า และมีประสิทธิภาพมากกว่าแบบอื่น อย่างไรก็ตามประสิทธิภาพการลดมิติข้อมูลของ PAA นี้เป็นวิธีที่เหมาะสมกับข้อมูลที่ผู้วิจัยได้เลือกมาทำนั้นในอนาคตการวิจัยครั้งต่อไปควรทำการทดสอบกับข้อมูลที่มีความนิยม เช่น ข้อมูล UCI dataset และใช้การเลือกคุณสมบัติแบบอื่นมาวัดประสิทธิภาพเป็นต้น

เอกสารอ้างอิง

[1] B. Xie, *et al.*, "Social Behavior Association and Influence in Social Networks," vol. 9, pp. 434-439, 2009.

- [2] (2011, 20 Dec.). *May 2011 - Mobile : SOCIAL NETWORKING, MOBILE, PRIVACY*. Available: <http://pewinternet.org/Shared-Content/Data-Sets/2011/May-2011--Mobile.aspx>
- [3] G. M. Papadourakis, *et al.*, "Clustering Analysis of Questionnaire for Ph.D. studies in Electrical and Information Engineering in Europe," Joseph Fourier, France, 2008.
- [4] Z. Chen, *et al.*, "An Improved Apriori Algorithm Based on Pruning Optimization and Transaction Reduction," vol. 11, 2011.
- [5] T. Hoya, "Graph Theoretic Techniques for Pruning Data and Their Applications," *IEEE TRANSACTIONS ON SIGNAL PROCESSING*, vol. 46, pp. 2574-2579, SEPTEMBER 1998 1998.
- [6] X. Cai-quan, *et al.*, "Association Rules Mining for Discussion Information in Group Deliberation Support System," *COF'11puter Science & Education (ICCSE 2013)*, vol. 13, pp. 67-70, April 26-28, 2013 2013.
- [7] กฤษฎากร กิ่งอุบล, "การใช้เกณฑ์ความต่างลำดับในการปรับปรุงกฎความสัมพันธ์จำแนกประเภทข้อมูล," *วิศวกรรมศาสตรมหาบัณฑิต, วิศวกรรมคอมพิวเตอร์, มหาวิทยาลัยเกษตรศาสตร์*, 2552.
- [8] W. Buathong, *et al.*, "Effect of data dimension reduction Analysis and Data Classification Performance of Decision Tree, Support Vector Machine and Naïve Bayes " presented at the The Eighth National Conference on Computing and Information Technology, 2012.
- [9] M. A. Hall, "Correlation-based Feature Selection for Machine Learning," Doctor of Philosophy, Department of Computer Science, University of Waikato, Hamilton, New Zealand, 1999.
- [10] ภรณ์ยาอามฤครัตน์และคณะ, "A Comparative Efficiency of Feature Selection and Neural Network Classification," presented at the The 5th National Conference on Computing and Information Technology, 2009.
- [11] V. Nuijian and P. Meesad, "A Comparison of Filter and Wrapper Approaches with Text Mining for Text Classification," *The Journal of Industrial Technology*, vol. 9, 2013.
- [12] H. Shi and W. Li, "Artificial neural networks with ant colony optimization for assessing performance of residential buildings," presented at the BioMedical Information Engineering, 2009.
- [13] P. Pengsiri and S. Pinitkan, "Predicting Residential Requirement from Personal Data by Data Classification

Techniques," in *The 6th National Conference on Information Technology : NCIT2014*, Nakhon Ratchasima, Thailand, 2014, pp. 550 - 555.

การวิเคราะห์แนวโน้มการใช้เทคโนโลยีสารสนเทศและการสื่อสารของประชากร ในประเทศไทยโดยใช้เหมืองข้อมูลร่วมกับการวิเคราะห์ค่าน้ำหนักองค์ประกอบ Trend Analysis Model for Information and Communication Technology Regarding the Thai Population Using Data Mining with Factor Loading

ณัฐวัลย์ กมขำ (Natwalai Komkham)¹ และผุสดี บุญรอด (Pudsadee Boonrawd)²

^{1,2}ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

natwalai.k@nso.go.th¹, pudsadee@kmutnb.ac.th²

บทคัดย่อ

งานวิจัยนี้ได้นำเสนอการวิเคราะห์แนวโน้มการใช้เทคโนโลยีสารสนเทศและการสื่อสารของประชากรในประเทศไทยโดยนำข้อมูลการสำรวจ ตั้งแต่ปี พ.ศ. 2552 – 2556 มาวิเคราะห์โดยใช้เหมืองข้อมูลด้วยวิธีการแบ่งประเภทข้อมูลออกเป็นกลุ่มอายุ รายได้ และค่าน้ำหนักความสัมพันธ์ของปัจจัยทั้งหมดที่ส่งผลต่อการใช้เทคโนโลยีสารสนเทศ ร่วมกับการวิเคราะห์ค่าน้ำหนักองค์ประกอบของกลุ่มปัจจัยในการยืนยันความสัมพันธ์ของปัจจัยนั้น ๆ เพื่อให้สามารถวิเคราะห์หรือเปรียบเทียบแนวโน้มการใช้เทคโนโลยีสารสนเทศและการสื่อสารในภาพรวม ซึ่งแสดงผลเป็นแบบกราฟหรือแผนภูมิ เพื่อให้ง่ายต่อการทำความเข้าใจ โดยมีค่าความเชื่อมั่นเท่ากับ 0.87 และมีผลการวิเคราะห์ค่าน้ำหนักองค์ประกอบของกลุ่มปัจจัย ตั้งแต่ 0.83 ถึง 0.97 ดังนั้นสรุปได้ว่าหน่วยงานภาครัฐและภาคเอกชนสามารถนำข้อมูลสารสนเทศไปช่วยวางแผนการดำเนินงาน การพัฒนาส่งเสริมและสนับสนุนการตัดสินใจการใช้เทคโนโลยีให้กระจายสู่ประชาชนได้ถูกต้อง และรวดเร็วมากยิ่งขึ้น

คำสำคัญ: เหมืองข้อมูล กฎความสัมพันธ์ การจัดแบ่งประเภทข้อมูล ค่าน้ำหนักองค์ประกอบ

Abstract

This research presents a trend analysis model for information and communication technology regarding the Thai population using data mining with factor loading. This was achieved by using data mining with age

and income classification and association rules for finding factor that affect to the use of information technology from the survey data during 2008-2013 with analysis of factor loading to compare trends and analyze the Thai population use of information and communication technology. Results were then depicted via graphs and charts to increase ease of understanding. According to the results, confidence at 0.87 with factor loading at 0.89 to 0.97. In conclusion, the results from this research can be used to help governments or businesses plan operations by increasing support to decision making of management teams with high accuracy.

Keywords: Data Mining, Association Rules, Data Classification, Factor Loading

1. บทนำ

ในภาวะสังคมปัจจุบันประเทศไทยได้ก้าวเข้าสู่ยุคของทางด่วนสารสนเทศ (Information Superhighway) เกิดจากการนำเอาเทคโนโลยีคอมพิวเตอร์และเทคโนโลยีโทรคมนาคม มาใช้งานร่วมกัน ซึ่งได้เข้ามามีบทบาทในชีวิตประจำวันของประชาชนเป็นอย่างมาก โดยแนวโน้มในอนาคตจะมีการใช้เทคโนโลยีสารสนเทศกันมากขึ้น [1] กระทรวงเทคโนโลยีสารสนเทศและการสื่อสารมีหน้าที่กำกับดูแล วางแผน ส่งเสริมพัฒนาและดำเนินกิจการเกี่ยวกับการใช้เทคโนโลยีสารสนเทศและการสื่อสาร โดยใช้ข้อมูลการสำรวจจากครัวเรือนในประเทศไทยจากสำนักงานสถิติแห่งชาติ เพื่อให้ทราบข้อมูลการใช้เทคโนโลยีสารสนเทศและการสื่อสารของประชากร

แต่ข้อมูลจากการสำรวจนั้นเป็นข้อมูลเชิงปริมาณที่เป็นรูปแบบกระจายทำให้ยากต่อการค้นหา วิเคราะห์และเปรียบเทียบผลการสำรวจในภาพรวม ซึ่งข้อมูลที่มีอยู่นั้นไม่ได้นำมาทำการวิเคราะห์หรือวัดผลข้อมูลการใช้เทคโนโลยีสารสนเทศ

ดังนั้นคณะผู้วิจัยจึงนำข้อมูลการสำรวจมาวิเคราะห์โดยใช้เหมืองข้อมูลด้วยวิธีการแบ่งกลุ่มข้อมูล เพื่อทำการจำแนกประเภทกลุ่มข้อมูลอายุ รายได้ และหากดูความสัมพันธ์ของข้อมูลการสำรวจ ที่ส่งผลกระทบต่อการใช้เทคโนโลยีสารสนเทศและการสื่อสารร่วมกับการวิเคราะห์ค่าน้ำหนักองค์ประกอบของกลุ่มปัจจัย ซึ่งใช้ในการอธิบายโครงสร้างความสัมพันธ์ระหว่างปัจจัยนั้น ๆ เพื่อให้หน่วยงานภาครัฐสามารถนำข้อมูลสารสนเทศไปวางแผนการดำเนินงานหรือสนับสนุนการตัดสินใจในการส่งเสริมพัฒนาได้ รวมทั้งช่วยให้ภาคเอกชนสามารถนำไปวางแผน กำหนดการส่งเสริมการตลาด และเล็งเห็นโอกาสการเติบโตของการใช้อุปกรณ์ด้านเทคโนโลยีได้ ซึ่งงานวิจัยนี้แบ่งออกเป็น 5 ส่วนคือ บทนำ ทฤษฎีและงานวิจัยที่เกี่ยวข้อง วิธีการดำเนินการวิจัย ผลการดำเนินงานและสรุปผลการวิจัย และข้อเสนอแนะ เนื้อหาต่อไปคือทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

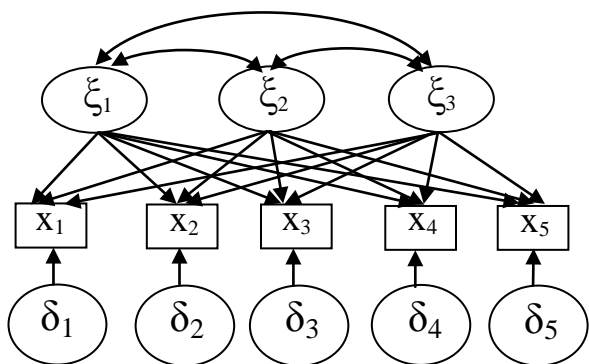
2.1 เหมืองข้อมูล (Data Mining)

เหมืองข้อมูลเป็นกระบวนการคัดกรองสารสนเทศที่มีอยู่ในฐานข้อมูลขนาดใหญ่ เพื่อนำข้อความรู้ที่ได้ไปใช้ประโยชน์ในการตัดสินใจ การทำเหมืองข้อมูลประกอบขึ้นด้วยการนำกระบวนการทางสถิติและการเรียนรู้ผ่านระบบคอมพิวเตอร์เพื่อสร้างตัวแบบ กฎเกณฑ์ รูปแบบ การพยากรณ์ และข้อความรู้จากฐานข้อมูลขนาดใหญ่ [2] สารสนเทศที่ได้อาจนำมาทำการจัดกลุ่ม (Clustering) [3] ซึ่งเป็นวิธีการกำหนดกลุ่มหรือประเภทที่มีความแตกต่างกันจำนวนหนึ่ง หรือการค้นหากฎความสัมพันธ์ (Association Rules) [4] โดยจะใช้การค้นหากฎความสัมพันธ์ในการหาความสัมพันธ์ของข้อมูลสองชุดหรือมากกว่าสองชุดขึ้นไปภายในกลุ่มข้อมูลที่มีขนาดใหญ่ในการหากฎความสัมพันธ์นั้นจะมีขั้นตอนวิธีการหาหลายวิธี แต่ขั้นตอนวิธีที่เป็นที่รู้จักและใช้อย่างแพร่หลาย คือ ขั้นตอนวิธีเอปพรอริ (Apriori) [5] นอกจากนี้ยังมีวิธีการสร้างตัวแบบการ

พยากรณ์และการจำแนก ที่ใช้สาระที่มีอยู่ในฐานข้อมูลซึ่งเป็นข้อมูลที่เกิดแล้วมาสร้างตัวแบบเพื่อคาดการณ์สิ่งที่จะเกิดในอนาคต การสร้างตัวแบบอาจจะใช้วิธีการจัดแบ่งประเภทข้อมูล (Data Classification) เป็นกระบวนการสร้างโมเดลจัดการข้อมูลให้อยู่ในกลุ่มข้อมูลที่แสดงให้เห็นความแตกต่างระหว่าง Class หรือ กลุ่มของข้อมูลและเพื่อทำนายว่า ข้อมูลนี้ควรจัดอยู่ใน Class ไດ [6] หรือการพยากรณ์ (Prediction) ซึ่งเป็นการกำหนดค่าที่คาดหวังว่าจะเกิดขึ้นของตัวแปรที่สนใจโดยอาศัยรูปแบบที่ได้จากข้อมูลที่มีอยู่เดิม อีกวิธีการหนึ่งในการทำเหมืองข้อมูลคือการค้นหารูปแบบที่เกิดขึ้นบ่อยหรือเกิดตามกันมาจากข้อมูลที่จัดเรียงเป็นลำดับ (Sequence) โดยจะเกี่ยวข้องกับเหตุการณ์ที่สัมพันธ์กับการเกิดที่เรียงลำดับกัน [7]

2.2 การวิเคราะห์องค์ประกอบ (Factor Analysis)

การวิเคราะห์องค์ประกอบ หรือบางครั้งเรียกว่าการวิเคราะห์ตัวประกอบ เป็นวิธีการทางสถิติที่พยายามจัดกลุ่มตัวแปรที่มีความสัมพันธ์กันให้อยู่ในกลุ่มเดียวกัน [8] ทำให้ทราบโครงสร้างและแบบของข้อมูลและหาองค์ประกอบร่วมของตัวแปรได้ เนื่องจากวิธีการวิเคราะห์องค์ประกอบจะหาค่าสัมประสิทธิ์สหสัมพันธ์ของตัวแปรทีละคู่แล้วรวมตัวแปรที่สัมพันธ์กันมากไว้ในปัจจัยเดียวกัน ซึ่งสังเกตจากการจับกลุ่มของตัวแปรหรือค่าสัมประสิทธิ์ความสัมพันธ์ระหว่างตัวแปร เช่น ถ้ามีตัวแปร 10 ตัว และตัวแปรนั้นมีความสัมพันธ์กันแบ่งออกได้ 10 กลุ่ม ในแต่ละกลุ่มตัวแปรมีความสัมพันธ์กันสูงที่เป็นเช่นนี้เพราะตัวแปรเหล่านั้นมีองค์ประกอบร่วมกัน ถ้าพบว่าองค์ประกอบร่วมของตัวแปรเหล่านั้นมีความสัมพันธ์กันสูง ดังนั้นการวิเคราะห์ข้อมูลต่าง ๆ อาจใช้องค์ประกอบร่วม 10 ตัวนั้นแทนตัวแปร 100 ตัว เป็นการลดจำนวนตัวแปรโดยการรวมตัวแปรหลาย ๆ ตัวให้อยู่ในปัจจัยเดียวกัน ปัจจัยที่ได้ถือว่าเป็นตัวแปรใหม่ ที่สามารถหาค่าข้อมูลของปัจจัยที่สร้างขึ้นได้ จึงสามารถนำปัจจัยดังกล่าวไปเป็นตัวแปรสำหรับการวิเคราะห์ต่อไป ทำให้สะดวกและประหยัดเวลาในการวิเคราะห์ข้อมูล การแยกว่าตัวแปรใดควรอยู่ในองค์ประกอบใดจะพิจารณาจากค่าน้ำหนักองค์ประกอบ (Factor Loading) ของตัวแปรต่าง ๆ ว่ามีค่ามากที่สุดอยู่ที่องค์ประกอบใด ก็จัดให้อยู่ในองค์ประกอบนั้น แต่มีข้อแม้ว่าค่าน้ำหนักองค์ประกอบควรมีค่าตั้งแต่ 0.6 ขึ้นไป โดยมีโครงสร้างการวิเคราะห์ ดังภาพที่ 1



ภาพที่ 1 : โครงสร้างการวิเคราะห์ห้อยค์ประกอบ

โดยที่ ξ คือ ตัวแปรแฝงหรือองค์ประกอบ

X คือ ตัวแปรที่สังเกตได้

δ คือ ตัวแปรความคลาดเคลื่อนหรือตัวแปรเฉพาะ

2.3 งานวิจัยที่เกี่ยวข้อง

จากการค้นคว้าเอกสารหรืองานวิจัยที่เกี่ยวข้อง พบว่ามีงานวิจัยที่เกี่ยวกับการศึกษาความต้องการและพฤติกรรมการใช้เทคโนโลยีสารสนเทศและการสื่อสาร โดยใช้วิธีการนำข้อมูลที่ได้ออกจากการตอบแบบสอบถามของกลุ่มตัวอย่างมาวิเคราะห์ผล เช่น การวิเคราะห์ความเห็นของการใช้เทคโนโลยีสารสนเทศและการสื่อสารพบว่าเพศมีผลระดับความเห็นเกี่ยวกับการใช้เทคโนโลยีสารสนเทศและการสื่อสารแตกต่างกันน้อยมากแต่อายุและการศึกษานั้นมีผลต่อระดับความเห็นที่แตกต่างกัน [9] หรือความถี่ในการใช้สารสนเทศว่าใช้เทคโนโลยีสารสนเทศมากน้อยเพียงใด [10] จะเห็นว่าการศึกษาพฤติกรรมการใช้เทคโนโลยีสารสนเทศด้วยวิธีการใช้แบบสอบถามเพียงอย่างเดียว จะทำให้ทราบแต่ข้อมูลในแบบปริมาณเท่านั้น ซึ่งถ้านำระบบสารสนเทศมาประยุกต์ใช้ในการวิเคราะห์ข้อมูลจะสามารถช่วยในการวิเคราะห์รวบรวมและนำเสนอข้อมูลได้มีประสิทธิภาพมากขึ้น

ส่วนของการนำเทคโนโลยีสารสนเทศมาประยุกต์ใช้เพื่อการวิเคราะห์ข้อมูลนั้น คณะผู้วิจัยได้ทำการศึกษาทฤษฎีของเหมืองข้อมูลในการหาความสัมพันธ์ [11] และการแบ่งกลุ่มข้อมูล [12],[13] เพื่อค้นหาความสัมพันธ์และรูปแบบที่มีอยู่จากข้อมูลจำนวนมาก คณะผู้วิจัยจึงมีแนวคิดในการค้นหาความสัมพันธ์ร่วมกับการจัดแบ่งประเภทข้อมูลมาประยุกต์ใช้เพื่อค้นหา วิเคราะห์ปัจจัยที่ส่งผลต่อการใช้เทคโนโลยี

สารสนเทศ หรือปัจจัยที่มีความสัมพันธ์กันซ่อนอยู่ และจำแนกประเภทข้อมูล เช่น กลุ่มอายุ กลุ่มรายได้ เป็นต้น เพื่อให้สามารถวิเคราะห์และเปรียบเทียบข้อมูลในภาพรวมได้

3. วิธีการดำเนินงานวิจัย

ในการวิเคราะห์แนวโน้มการใช้เทคโนโลยีสารสนเทศและการสื่อสารของประชากรในประเทศไทยโดยใช้เหมืองข้อมูลร่วมกับการวิเคราะห์ค่าน้ำหนักองค์ประกอบ เพื่อให้สามารถวิเคราะห์หรือเปรียบเทียบแนวโน้มการใช้เทคโนโลยีสารสนเทศและการสื่อสารในภาพรวม ตั้งแต่ปี พ.ศ.2552 – 2556 และแสดงผลเป็นแบบกราฟหรือแผนภูมิเพื่อให้เข้าใจง่าย และบรรลุตามวัตถุประสงค์ โดยแบ่งการพัฒนาออกเป็น 2 ส่วนด้วยกัน ได้แก่

3.1 การวิเคราะห์ปัญหาของข้อมูลผลการสำรวจ

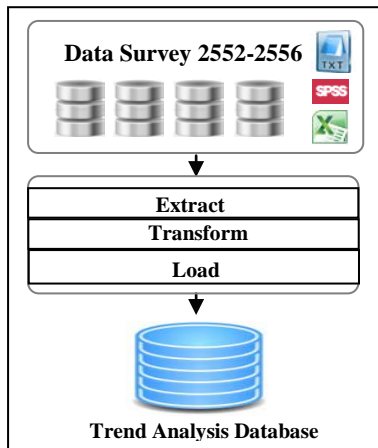
กระบวนการในการรายงานผลการจัดทำข้อมูลสถิติจากการสำรวจการใช้เทคโนโลยีสารสนเทศและการสื่อสารของประชากรในประเทศไทยของสำนักงานสถิติแห่งชาติ ได้จากการประมวลผล สรุปผลผ่านโปรแกรม SPSS ซึ่งรายงานผลการสำรวจจะอยู่ในรูปแบบของกระดาษ โดยนำเสนอข้อมูลแบบเชิงปริมาณ และกราฟข้อมูลจำนวนหรือร้อยละของข้อมูลสำรวจได้ในแต่ละปีเท่านั้น ซึ่งไม่มีระบบสารสนเทศเข้ามาช่วยในการประมวลผลหรือวิเคราะห์ผลและเปรียบเทียบแนวโน้มของแต่ละปีหรือหาความสัมพันธ์ของปัจจัยต่าง ๆ

3.2 กระบวนการวิเคราะห์แนวโน้มการใช้เทคโนโลยีสารสนเทศ และการสื่อสาร

การวิเคราะห์ข้อมูลและปัจจัยที่มีผลต่อการใช้เทคโนโลยีสารสนเทศและการสื่อสาร โดยแบ่งการวิเคราะห์ออกเป็น 3 ส่วนคือ การเตรียมข้อมูล การวิเคราะห์และการแสดงผล

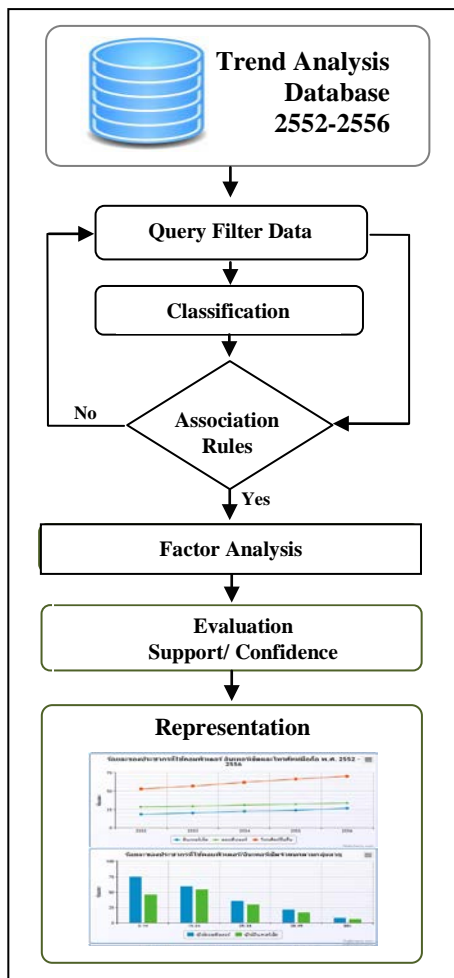
ส่วนที่ 1 การเตรียมข้อมูล โดยนำข้อมูลการสำรวจการใช้เทคโนโลยีสารสนเทศและการสื่อสารตั้งแต่ปี พ.ศ. 2552 - 2556 โดยข้อมูลการสำรวจเหล่านี้จะนำมาทำการดึงข้อมูล (Extract) จากแหล่งข้อมูลที่ต่างกัน ซึ่งเป็นข้อมูลจากไฟล์ excel ไฟล์ SPSS และไฟล์ Text โดยเลือกใช้เฉพาะข้อมูลที่เกี่ยวข้องกับสิ่งที่ทำการวิเคราะห์ แล้วทำการจัดรูปแบบข้อมูล (Transform) เพื่อเตรียมข้อมูลให้อยู่ในรูปแบบที่สอดคล้องกัน

ตรวจสอบและแก้ไขข้อมูลให้ถูกต้อง จากนั้นนำข้อมูลที่ผ่านการจัดรูปแบบแล้วไปเก็บไว้ในฐานข้อมูล (Load) ดังภาพที่ 2



ภาพที่ 2 : แผนภาพกระบวนการเตรียมข้อมูล

ส่วนที่ 2 การวิเคราะห์ โดยคณะผู้วิจัยได้ใช้โปรแกรม WEKA ในการประมวลผลข้อมูล ซึ่งมีขั้นตอนดังภาพที่ 3



ภาพที่ 3 : โครงสร้างการวิเคราะห์แนวโน้มการใช้เทคโนโลยีสารสนเทศและการสื่อสาร

จากภาพที่ 3 ทำการวิเคราะห์ข้อมูล โดยนำข้อมูลผ่านกระบวนการคัดกรองข้อมูลที่เกี่ยวข้อง แล้วนำมาจำแนกประเภทข้อมูลเช่น กลุ่มอายุ กลุ่มรายได้ โดยมีตัวอย่างการจำแนกข้อมูล ดังภาพที่ 4

```

=== Confusion Matrix ===
      a   b   c   d   e  <-- classified as
54966  11  20  48  43 | a = 35-49
 17 29767  15  25  19 | b = 25-34
 28  15 26955  22  37 | c = 15-24
 32  15  26 62199  42 | d = 50
 19  5  10  14 48242 | e = 6-14
    
```

ภาพที่ 4 : ผลการแบ่งกลุ่มอายุ

ตัวอย่างชุดคำสั่งในการเรียกใช้กลุ่มอายุที่ได้จากการจำแนกกลุ่ม โดยนำไปใช้ในการคัดเลือกข้อมูล เพื่อไปแสดงผลข้อมูลทั้งแบบแผนภูมิ กราฟ และข้อมูลเชิงตาราง ดังภาพที่ 5

```

1 SELECT
2 com_2555.REG,com_2555.REG1,reg.REG_E,
3 com_2555.CWT,com_2555.CWT1,cwt.CWT_E,com_2555.AREA,
4 com_2555.AREA1,area.AREA_E,
5 com_2555.SEX,com_2555.AGE,
6 if(com_2555.AGE between 1 and 14 , '6-14',
7 if(com_2555.AGE between 15 and 24 , '15-24',
8 if(com_2555.AGE between 25 and 34 , '25-34',
9 if(com_2555.AGE between 35 and 49 , '35-49',
10 if(com_2555.AGE >= 50 , '50+', '-')))) AS ages,
11 com_2555.MARITAL,com_2555.MERITAL,
12 com_2555.MERITAL1,com_2555.MERITAL_E,com_2555.GRADE,
13 com_2555.GRADE1,grade.GRADE,
14 com_2555.SUBJECT,com_2555.SUBJECT1,
15 com_2555.EDU_ID,com_2555.EDU,
16 com_2555.EDU_E,com_2555.WK_12MONTH,
17 com_2555.STATUS,com_2555.STATUS1,com_2555.STATUS_E,
18 com_2555.STATUS1,com_2555.STATUS1,com_2555.STATUS_E
19 FROM
20 com_2555
21 INNER JOIN reg ON com_2555.REG = reg.REG_ID
22 INNER JOIN cwt ON com_2555.CWT = cwt.CWT_ID
23 INNER JOIN area ON com_2555.AREA = area.AREA_ID
24 INNER JOIN merital ON com_2555.MARITAL = merital.MERITAL_ID
25 INNER JOIN grade ON com_2555.GRADE = grade.GRADE_ID
26 INNER JOIN edu ON com_2555.EDU_ID = edu.EDU_ID
27 INNER JOIN status_wk ON com_2555.STATUS = status_wk.STATUS_ID
28
    
```

ภาพที่ 5 : แผนภาพชุดคำสั่งในการเรียกใช้ของข้อมูลการแบ่งกลุ่มอายุ

จากนั้นทำการหาความสัมพันธ์ข้อมูลที่มีผลกับการใช้เทคโนโลยี โดยใช้โปรแกรม WEKA ในการประมวลผลข้อมูล ใช้ขั้นตอนวิธีแบบเอปปริออริ (Apriori) โดยมีตัวอย่างกฎความสัมพันธ์ คือ ถ้าเป็นผู้ที่มีอายุ 25-34 ปี เป็นผู้มีงานทำ จบการศึกษาระดับปริญญาตรี ประกอบอาชีพเป็นลูกจ้างรัฐบาล และมีรายได้ตั้งแต่ 20,000-24,900 บาท แล้วจะใช้คอมพิวเตอร์ ดังภาพที่ 6

559. WK=Yes CAREER=Government_Employee 26630 ==> INCOME=20K-24.9K USED_COM=Yes EDU=B.D. 26630 <conf(1)> lift(1)
560. CAREER=Government_Employee 26630 ==>WK=Yes INCOME=20K-24.9K USED_COM=Yes EDU=B.D. 26630 <conf(1)> lift(1)
561. WK=Yes CAREER=Government_Employee INCOME=20K-24.9K USED_COM=Yes EDU=B.D. 26630 ==>AGE=25-34 26630 <conf(1)> lift(1)
562. AGE=25-34 CAREER=Government_Employee INCOME=20K-24.9K USED_COM=Yes EDU=B.D. 26630 ==> WK=Yes 26630 <conf(1)> lift(1)
563. AGE=25-34 WK=Yes CAREER=Government_Employee USED_COM=Yes EDU=B.D. 26630 ==> INCOME=20K-24.9K 26630 <conf(1)> lift(1)
564. AGE=25-34 WK=Yes CAREER=Government_Employee INCOME=20K-24.9K EDU=B.D. 26630 ==> USED_COM=Yes 26630 <conf(1)> lift(1)
565. AGE=25-34 WK=Yes CAREER=Government_Employee INCOME=20K-24.9K USED_COM=Yes 26630 ==> EDU=B.D. 26630 <conf(1)> lift(1)
566. CAREER=Government_Employee INCOME=20K-24.9K USED_COM=Yes EDU=B.D. 26630 ==> AGE=25-34 WK=Yes 26630 <conf(1)> lift(1)
567. WK=Yes CAREER=Government_Employee USED_COM=Yes EDU=B.D. 26630 ==> AGE=25-34 INCOME=20K-24.9K 26630 <conf(1)> lift(1)
568. WK=Yes CAREER=Government_Employee INCOME=20K-24.9K EDU=B.D. 26630 ==> AGE=25-34 USED_COM=Yes 26630 <conf(1)> lift(1)
569. WK=Yes CAREER=Government_Employee INCOME=20K-24.9K USED_COM=Yes 26630 ==> AGE=25-34 EDU=B.D. 26630 <conf(1)> lift(1)
570. AGE=25-34 CAREER=Government_Employee USED_COM=Yes EDU=B.D. 26630 ==> WK=Yes INCOME=20K-24.9K 26630 <conf(1)> lift(1)
571. AGE=25-34 CAREER=Government_Employee INCOME=20K-24.9K EDU=B.D. 26630 ==> WK=Yes USED_COM=Yes 26630 <conf(1)> lift(1)

ภาพที่ 6 : ตัวอย่างผลการหาความสัมพันธ์

ซึ่งกฎที่ได้ต้องผ่านค่าสนับสนุน (Support) และค่าความเชื่อมั่น (Confidence) โดยมีสูตรการคำนวณดังสมการที่ 1 และสมการที่ 2

1) Support Factor เป็นค่าที่บ่งบอกว่าเหตุการณ์ A กับ B มีความถี่ในการเกิดขึ้นมากน้อยแค่ไหน

$$A \rightarrow B : \text{Support Factor} = AUB \quad (1)$$

2) Confident Factor เป็นค่าที่บอกว่า เมื่อเกิดเหตุการณ์ B แล้ว มีโอกาสที่จะเกิดเหตุการณ์ A มากน้อยแค่ไหน

$$A \rightarrow B : \text{Confident Factor} = P(A|B) \quad (2)$$

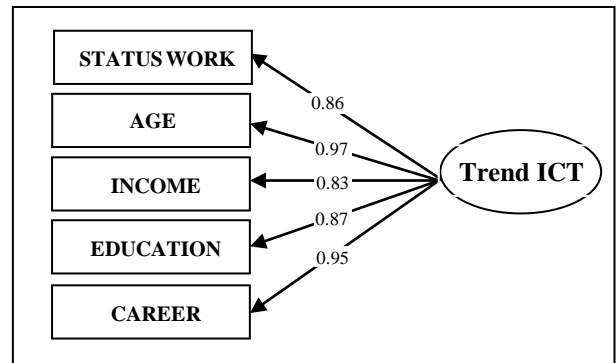
โดยมีค่าสนับสนุนเท่ากับ 0.1 และค่าความเชื่อมั่นมากกว่า 0.8 จากนั้นทำการวิเคราะห์องค์ประกอบของกลุ่มปัจจัยจากค่าน้ำหนักองค์ประกอบที่มีความสัมพันธ์กันระหว่างกันของข้อมูล เพื่อให้ทราบโครงสร้างแบบแผนของข้อมูลตลอดจนค่าน้ำหนักแต่ละตัวประกอบ โดยใช้โปรแกรม LISREL ในการประมวลผลข้อมูล

ส่วนที่ 3 การแสดงผล นำเอากลุ่มปัจจัยและผลการวิเคราะห์ที่หาความสัมพันธ์ ที่ได้จากส่วนที่ 2 มาแสดงผลแบบกราฟหรือแผนภูมิ เพื่อให้ง่ายต่อการทำความเข้าใจ

4. ผลการวิเคราะห์แนวโน้มการใช้เทคโนโลยีสารสนเทศและการสื่อสาร

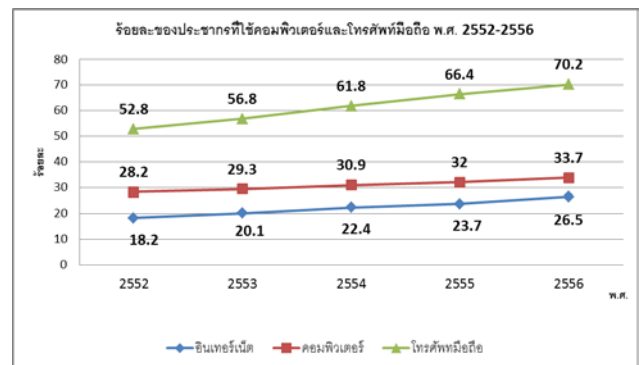
จากการวิเคราะห์แนวโน้มการใช้เทคโนโลยีสารสนเทศและการสื่อสารของประชากรในประเทศไทยโดยใช้เหมืองข้อมูลร่วมกับการวิเคราะห์ค่าน้ำหนักองค์ประกอบ ทำให้ทราบกลุ่มปัจจัยที่มีผลต่อการใช้เทคโนโลยีสารสนเทศและการ

สื่อสารซึ่งแบ่งออกเป็นปัจจัยที่มีผลกับการใช้คอมพิวเตอร์ อินเทอร์เน็ต และโทรศัพท์มือถือ ประกอบด้วย รายได้ สถานะการทำงาน ช่วงอายุ การศึกษา และอาชีพ ซึ่งมีค่าน้ำหนักของ รายได้ สถานะการทำงาน อายุ การศึกษา อาชีพ ซึ่งมีความสัมพันธ์ระหว่างปัจจัยในระดับสูง มีค่าตั้งแต่ 0.83 ถึง 0.97 โดยมีตัวอย่างโมเดลองค์ประกอบกลุ่มปัจจัย ดังภาพที่ 7



ภาพที่ 7 : แผนภาพโมเดลองค์ประกอบกลุ่มปัจจัย ที่ส่งผลต่อการใช้เทคโนโลยีสารสนเทศและการสื่อสาร

จากนั้นนำเอาผลการวิเคราะห์มาแสดงผลแบบกราฟและแผนภูมิ เพื่อให้ง่ายต่อการทำความเข้าใจเปรียบเทียบข้อมูลย้อนหลัง 5 ปี ดังภาพที่ 8



ภาพที่ 8 : แผนภาพการวิเคราะห์แนวโน้มการใช้เทคโนโลยี

5. สรุปผลการวิจัย

งานวิจัยนี้ได้เสนอ การวิเคราะห์แนวโน้มการใช้เทคโนโลยีสารสนเทศและการสื่อสารของประชากรในประเทศไทยโดยใช้เหมืองข้อมูลร่วมกับการวิเคราะห์ค่าน้ำหนักองค์ประกอบ จากผลการวิจัยพบว่า สามารถค้นหาปัจจัยหรือองค์ประกอบที่มีความสัมพันธ์กัน และส่งผลต่อการใช้เทคโนโลยีสารสนเทศ

และการสื่อสารจากปัจจัยที่อยู่มีมากมายในแบบสอบถาม แล้วทำการวิเคราะห์ห้วงค์ประกอบเพื่อยืนยันความสัมพันธ์ระหว่างปัจจัยว่าปัจจัยนั้น ๆ ส่งผลต่อการใช้เทคโนโลยีสารสนเทศ โดยมีค่าน้ำหนักองค์ประกอบของกลุ่มปัจจัยซึ่งมีความสัมพันธ์ระหว่างปัจจัยในระดับสูงโดยมีค่าตั้งแต่ 0.83 ถึง 0.97 และมีค่าระดับความเชื่อมั่นเท่ากับ 0.87 ซึ่งสามารถนำข้อมูลไปใช้ในการพัฒนาส่งเสริมและผลักดันให้ประชาชนมีความรู้ความเข้าใจเกี่ยวกับเทคโนโลยีสารสนเทศและการสื่อสาร เพื่อยกระดับการทำงาน เสริมสร้างความเท่าเทียมในสังคมและเพิ่มโอกาสในการเรียนรู้และพัฒนาคุณภาพชีวิตได้ทำทันกระแสเทคโนโลยีสารสนเทศในปัจจุบัน งานวิจัยในอนาคต ควรทำการวิเคราะห์เพื่อหาองค์ประกอบที่ส่งผลกับการใช้และไม่ใช้เทคโนโลยีสารสนเทศและการสื่อสารให้มีความสอดคล้องกับเทคโนโลยี ที่มีการเปลี่ยนแปลงไปในแต่ละปีโดยใช้วิธีโครงข่ายประสาทเทียม การวิเคราะห์การถดถอย หรือ วิธีซัพพอร์ตเวกเตอร์แมชชีนร่วมกับการพยากรณ์แนวโน้มการมี การใช้เทคโนโลยีสารสนเทศและการสื่อสารของประชากรในประเทศไทยต่อไปในอนาคต และเพื่อให้ได้กลุ่มหรือกลุ่มปัจจัยที่มีความถูกต้องและทันสมัยอยู่เสมอ

เอกสารอ้างอิง

- [1] สำนักงานสถิติแห่งชาติ.2555 . “สำรวจการมี การใช้เทคโนโลยีสารสนเทศและการสื่อสารในครัวเรือน พ.ศ. 2555,” กรุงเทพมหานคร: บริษัท เท็กซ์ แอนด์ เจอร์นัล พับลิเคชั่น จำกัด, 2556.
- [2] คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย. “หนึ่งทศวรรษไอทีจุฬา : การเพิ่มศักยภาพการจัดการด้วยเทคโนโลยีสารสนเทศ” กรุงเทพมหานคร: 2544.
- [3] He, Q., et al. “Clustering in extreme learning machine feature space”. *Neurocomputing*. Vol.128, pp. 88–95, 2014.
- [4] อนุศักดิ์ วงศ์มีแก้ว, “การค้นหากฎความสัมพันธ์ของการเข้าใช้งานเว็บไซต์,” การค้นคว้าอิสระ สาขาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยอุบลราชธานี, 2550.
- [5] อาทิตยาพร โรจรัตน์ และวรารัตน์ รุ่งวรวิฑู. “การค้นพบรูปแบบการวางแผนการลงทะเบียนเรียนของนักศึกษาโดยใช้การทำเหมืองความสัมพันธ์ กรณีศึกษาหลักสูตรเทคโนโลยีสารสนเทศและการสื่อสาร,” in the 9TH National Conference on Computing and Information Technology, 2013, pp. 147-153.
- [6] ภักดี ดาดิยะ, “ระบบช่วยตัดสินใจด้านการกำหนดแผนปฏิบัติงานสอบบัญชีสหกรณ์,” ปัญหาพิเศษ วิทยาศาสตร์มหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ บัณฑิตวิทยาลัยมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2552.
- [7] J. Han, M. Kamber, and J. Pei, *Data Mining Concepts and Techniques*, Morgan Kaufmann, pp. 383-403, 2006.
- [8] อัสมะ หะยิมอะหะซอและ, “การวิเคราะห์องค์ประกอบ,” วารสารวิศลปริทัศน์. ปีที่ 5. ฉบับที่ 5, ตุลาคม 2548.
- [9] สมศักดิ์ คงเทศ, “การใช้เทคโนโลยีสารสนเทศและการสื่อสาร ICT ของครูในสถานศึกษา ตำบลบางนายสี อำเภอดงแก้ว จังหวัดพังงา,” วิทยานิพนธ์ คุรุศาสตรมหาบัณฑิต สาขาวิชาการบริหารการศึกษา คณะครุศาสตร์ มหาวิทยาลัยราชภัฏสวนดุสิต, 2554.
- [10] เหมิกา กลิ่นขจร, “พฤติกรรมการใช้เทคโนโลยีสารสนเทศและการสื่อสารเพื่อการศึกษาของนักศึกษาระดับปริญญาบัณฑิต คณะมัณฑนศิลป์ มหาวิทยาลัยศิลปากร,” การค้นคว้าอิสระ ศึกษาศาสตร์มหาบัณฑิต ภาควิชาเทคโนโลยีการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศิลปากร, 2553.
- [11] Mei-Hua Hsu., “Proposing a charting recommender system for second-language nurses”. *Expert Systems with Applications*, Vol. 38, pp. 9281–9286, 2011.
- [12] Rashedur M. Rahman., “Using and comparing different decision tree classification techniques for mining ICDDR,B Hospital Surveillance data” *Expert Systems with Applications*, Vol. 38, pp. 11421-11436, 2011.
- [13] Peter C. Austin. “Using methods from the data-mining and machine-learning literature for disease classification and prediction: a case study examining classification of heart failure subtypes”. *Expert Systems with Application*, Vol. 66, pp. 398-407, 2013.

การเลือกปัจจัยที่ส่งผลต่อผลสัมฤทธิ์ทางการเรียน ตามระดับคะแนนของกลุ่มผู้เรียน ด้วย อัลกอริทึมการจำแนกกลุ่มบนปริภูมิย่อย

Selection Factors Affecting Learning Achievement Following Grade of Students Cluster by Subspace Clustering Algorithms

วีระยุทธ พิมพาภรณ์ (Werayut Pimpaporn)^{1,2} และ พยุง มีสัจ (Phayung Meesad)²

¹คณะวิทยาศาสตร์ ศรีราชา มหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตศรีราชา ชลบุรี

²คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ กรุงเทพมหานคร

บทคัดย่อ

งานวิจัยฉบับนี้มุ่งเน้นศึกษาวิธีการเลือกปัจจัยที่ส่งผลต่อผลสัมฤทธิ์ทางการเรียน ตามระดับคะแนนของกลุ่มผู้เรียน โดยใช้ชุดข้อมูลการทดสอบในลักษณะต่าง ๆ ระหว่างเรียน ซึ่งเป็นปัจจัยที่เกี่ยวข้องกับผลสัมฤทธิ์ทางการเรียน ทั้งนี้ กระบวนการเลือกปัจจัย ผู้วิจัยประยุกต์ใช้อัลกอริทึมการจำแนกกลุ่มบนปริภูมิย่อย (Subspace Clustering Algorithms : SCA) เพื่อใช้ในการจำแนกกลุ่มข้อมูล

จากผลการวิจัยพบว่าประสิทธิภาพโดยรวมของอัลกอริทึมให้ค่าความถูกต้องในการจำแนกกลุ่มอยู่ที่ 82.92% ทั้งนี้ปัจจัยที่สามารถจำแนกกลุ่มข้อมูลได้อย่างมีประสิทธิภาพได้แก่ ค่าคะแนนในกระบวนการเรียนออนไลน์ (Score.eLearning) เป็นปัจจัยของระดับคะแนน 5 เกรด ลำดับถัดมาได้แก่ค่าคะแนนสอบกลางภาค (Score.Mit) ค่าคะแนนสอบปลายภาค (Score.Final) และชั้นปีของผู้เรียน (Level) เป็นปัจจัยของระดับคะแนน 4 เกรด อีกทั้งระดับคะแนน A มีค่าประสิทธิภาพในการจำแนกกลุ่มสูงที่สุดที่ระดับ 100% โดยใช้ปัจจัย 4 ปัจจัยได้แก่ จำนวนครั้งในการสอบผ่านแบบทดสอบระหว่างเรียน (NumOfPass) แบบทดสอบประจำบทเรียนที่ 6 (FT-6) คะแนนสอบกลางภาค (Score.Mit) ค่าคะแนนสอบปลายภาค (Score.Final)

คำสำคัญ: จำแนกกลุ่มข้อมูล ปริภูมิย่อย ผลสัมฤทธิ์

Abstract

This research focused on the factors selecting affects to the achievement of student's grade by using the examination data set that is related with the

learning achievement. For the factors identify process, the researcher developed the Subspace Clustering Algorithms (SCA) in order to identify the cluster.

The result found that the accuracy in cluster of Subspace Clustering Algorithms (SCA) is at 82.92%. The effective factors that can cluster was the e-learning score (Score.eLearning) at 5 grade levels. Next, the factors that were the midterm score, the final score and the student's level were at 4 grade levels.

For the grade A, the 4 factors that were number of exam pass (NumOfPass), the chapter 6 of exam (FT-6), the midterm score (Score.Mit) and the final score (Score.Final) were effectively to cluster at 100%.

Keywords: Clustering, Subspace, Achievement

1. บทนำ

เทคโนโลยีการศึกษาในปัจจุบัน มีการปรับเปลี่ยนรูปแบบ และมีการพัฒนาอย่างต่อเนื่อง ซึ่งเป็นผลโดยตรงจากเทคโนโลยีคอมพิวเตอร์ที่เปลี่ยนแปลงไปอย่างรวดเร็ว ส่งผลทำให้องค์กรการศึกษาเริ่มตระหนักถึงการปรับเปลี่ยนรูปแบบการเรียนการสอน เพื่อเป็นการดึงข้อดีของเทคโนโลยีคอมพิวเตอร์ในปัจจุบันมาใช้ให้เกิดประโยชน์สูงสุดกับผู้เรียนและผู้สอน โดยเป็นจุดเน้นเทคนิคเหมืองข้อมูลถูกนำมาประยุกต์ใช้งานด้านการศึกษาอย่างแพร่หลาย

เทคนิคเหมืองข้อมูล (Data Mining) จึงถูกนำมาประยุกต์ใช้ในการพัฒนาระบบการจัดการเรียนการสอนในหลากหลายมิติ ทั้งการใช้เพื่อเป็นกระบวนการค้นกรอง

สารสนเทศจากข้อมูลที่เกิดขึ้นระหว่างการเรียนการสอน ใช้สำหรับทำนายแนวโน้มหรือพฤติกรรมของผู้เรียนที่จะเกิดขึ้นในอนาคตด้วยการวิเคราะห์ข้อมูลในอดีต รวมทั้งการสร้าองค์ความรู้ใหม่เพื่อใช้สำหรับตัดสินใจในการจัดการเรียนการสอนให้เกิดประสิทธิภาพสูงสุด

งานวิจัยฉบับนี้มุ่งเน้นศึกษาวิธีการเลือกปัจจัยที่ส่งผลโดยตรงต่อผู้เรียนเมื่อเข้าสู่กระบวนการเรียนการสอน ทั้งนี้ผู้เรียนทุกคนจะผ่านกระบวนการทดสอบในลักษณะต่าง ๆ เพื่อเก็บเป็นข้อมูลสำหรับวิเคราะห์ โดยกระบวนการเลือกปัจจัย ผู้วิจัยใช้อัลกอริทึมการจำแนกกลุ่มบนปริภูมิย่อย (Subspace Clustering Algorithms : SCA) ในการจำแนกกลุ่มข้อมูลผลการทดสอบของผู้เรียนตลอดภาคการศึกษา เพื่อใช้ในการเลือกปัจจัยที่ส่งผลกระทบต่อผลสัมฤทธิ์ทางการเรียน แบ่งแยกตามระดับคะแนนของกลุ่มผู้เรียน

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ผลสัมฤทธิ์ทางการเรียน (Learning Achievement)

ผลสัมฤทธิ์ทางการเรียน เป็นสิ่งที่สำคัญสำหรับการพิจารณาคุณภาพของการจัดการเรียนการสอน โดยเน้นที่การวัดความรู้ ความเข้าใจ และความสามารถของผู้เรียนที่เกิดจากการจัดการเรียนการสอนด้วยวิธีการทดสอบ เพื่อให้ได้ข้อมูลสำหรับนำไปเปรียบเทียบกับเกณฑ์มาตรฐานที่ได้กำหนดไว้ล่วงหน้า จากกระบวนการดังกล่าวผู้เรียนจะถูกแยกออกเป็นกลุ่มตามระดับความรู้ ในการวิจัยครั้งนี้ผู้วิจัยทำการศึกษาระดับเชิงพุทธิพิสัย (Cognitive Domain) ของกลุ่มผู้เรียนแต่ละกลุ่มแยกตามระดับความรู้ตามมาตรฐานเกณฑ์ระดับอุดมศึกษา แยกออกเป็น 8 ระดับคะแนน คือ A (Excellent), B+ (Very Good), B (Good), C (Average) จนถึงระดับ F ซึ่งถือเป็นคลาสคำตอบ (Target Class) ของชุดข้อมูล (Feature set)

2.2 การวิเคราะห์กลุ่ม (Clustering analysis)

การวิเคราะห์กลุ่มคือเทคนิคในการเรียนรู้แบบไม่มีผู้สอน (Unsupervised learning technique) ซึ่งมีเป้าหมายเพื่อการจำแนกกลุ่มข้อมูลที่มีคุณลักษณะคล้ายกันอยู่ในกลุ่มเดียวกัน [1] โดยข้อมูลแต่ละกลุ่มจะถูกเรียกว่า คลัสเตอร์ (Cluster) การวิเคราะห์หรือจำแนกกลุ่มข้อมูลนั้นสามารถแบ่งออกได้เป็น 2 ประเภทได้แก่ วิธีการแบบลำดับชั้น (Hierarchical

algorithms) และ วิธีการแบบไม่เป็นลำดับชั้น (Non-hierarchical algorithms) [2]

2.3 การจำแนกกลุ่มบนปริภูมิย่อย (Subspace Clustering)

การจัดกลุ่มบนปริภูมิย่อย (Subspace clustering) [3] ถือได้ว่าเป็นส่วนขยายของการจัดกลุ่มแบบดั้งเดิม (Traditional Clustering) [4] โดยมีวัตถุประสงค์หลักเพื่อการค้นหา กลุ่มข้อมูล (Clusters) ปริภูมิย่อยของชุดข้อมูล มักจะถูกนำเสนอในการค้นหา กลุ่มข้อมูลที่มีลักษณะ มิติสูง (High Dimensional Data) มิติข้อมูลไม่เกี่ยวข้องกัน (Dimensions Irrelevant) รวมถึงชุดข้อมูลที่มี ข้อมูลรบกวน (Noisy Data) ปะปนอยู่กับชุดข้อมูล

2.4 K-Means Clustering

K-means clustering [5] คือวิธีการจำแนกกลุ่มข้อมูลด้วยวิธีการแบ่งข้อมูลอัตโนมัติตามค่า k ที่กำหนด โดยกระบวนการทำงานเลือกค่า k เริ่มต้นสำหรับเป็นค่ากลางในการจัดกลุ่ม และปรับค่าตามกระบวนการดังนี้ 1) เลือกข้อมูล d_i สำหรับวัดระยะห่างกับค่า K เริ่มต้นทุกค่า 2) กำหนดชุดข้อมูล d_i ให้กับ K ที่ใกล้ที่สุด และปรับค่า K ใหม่ให้เป็นค่ากลางของกลุ่มข้อมูล และหยุดเมื่อค่า K ไม่เปลี่ยนแปลง [6] ดังนั้นการใช้ K-means clustering จึงสามารถนำมาใช้งานเมื่อทราบจำนวนกลุ่มที่ต้องการจำแนกที่แน่นอน

2.5 การวัดประสิทธิภาพแบบจำลอง

วิธีการวัดประสิทธิภาพการพยากรณ์กับชุดข้อมูลในการวิจัยครั้งนี้ใช้วัดประสิทธิภาพการพยากรณ์ข้อมูลตามแนวคิดด้านการค้นคืนสารสนเทศ (Information Retrieval) ด้วยค่าความถูกต้อง (Accuracy) ซึ่งเป็นการคำนวณจากตาราง Confusion Matrix [2], [7]

3. วิธีดำเนินการวิจัย

3.1 ศึกษาปัญหาและวิเคราะห์ข้อมูล

การวิจัยครั้งนี้เน้นการวิเคราะห์ข้อมูลที่ได้จากการทดสอบเพื่อวัดความรู้ความสามารถด้านพุทธิพิสัย (Cognitive Domain) ที่ส่งผลกระทบต่อผลสัมฤทธิ์ทางการเรียน ดังนั้นผู้วิจัยจึง

ใช้แบบทดสอบวัดความรู้ ความเข้าใจ ที่สอดคล้องกับวัตถุประสงค์และเป้าหมายของรายวิชา

ข้อมูลที่ใช้สำหรับวิเคราะห์ผู้วิจัยแบ่งข้อมูลออกเป็น 4 ส่วนดังนี้คือ 1. ข้อมูลที่ได้จากการทดสอบเพื่อวัดความรู้ความสามารถด้านพุทธิพิสัย (Cognitive Domain) ที่ส่งผลต่อผลสัมฤทธิ์ทางการเรียน ดังนั้นผู้วิจัยจึงใช้แบบทดสอบวัดความรู้ ความเข้าใจ ที่สอดคล้องกับวัตถุประสงค์และเป้าหมายของรายวิชา โดยแบ่งการทดสอบออกเป็น 3 ประเภทดังนี้ 1) แบบทดสอบก่อนเรียน (Pre-test) เป็นแบบทดสอบที่ใช้ประเมินผู้เรียนก่อนดำเนินการเรียนการสอน เพื่อสำรวจความพร้อมของผู้เรียน และวัดความรู้พื้นฐานเดิมของผู้เรียน 2) แบบทดสอบประจำบทเรียน (Formative test) เป็นการทดสอบตามวัตถุประสงค์การเรียนรู้ของแต่ละหน่วยการเรียนรู้ เพื่อสำรวจความรู้ ความเข้าใจ ที่ผู้เรียนได้จากการเรียนผ่านระบบการเรียนแบบผสมผสาน โดยแบ่งออกเป็น 8 หน่วยการเรียนรู้ และ 3) แบบทดสอบหลังเรียน (Post-test) เป็นแบบทดสอบที่ใช้ในการประเมินผลการเรียนของผู้เรียนเมื่อสิ้นสุดการเรียนทั้งรายวิชา การทดสอบทั้ง 3 ส่วนคิดค่าคะแนนเป็น 15% ของรายวิชา 2. คะแนนการทดสอบประจำภาคการศึกษา โดยแบ่งออกเป็น 2 ส่วนดังนี้ 1) การทดสอบกลางภาค (Midterm Examination) ครอบคลุมเนื้อหาในหน่วยการเรียนรู้ที่ 1 - 4 คิดเป็นค่าคะแนน 30% ของรายวิชา และ 2) การทดสอบปลายภาคเรียน (Final Examination) ครอบคลุมเนื้อหาในหน่วยการเรียนรู้ที่ 5 - 8 คิดเป็นค่าคะแนน 30% ของรายวิชา 3. ค่าคะแนนการฝึกปฏิบัติ คิดเป็น 25% ของคะแนนในรายวิชา และ 4. ข้อมูลทั่วไปของผู้เรียน ชุดข้อมูลในส่วนสุดท้ายได้แก่ข้อมูลทั่วไปของผู้เรียน เช่น ข้อมูลภาคการศึกษา ชั้นปีที่เรียน เป็นต้น

3.2 การเตรียมข้อมูล (Data pre-processing stage)

งานวิจัยขั้นนี้ผู้วิจัยเลือกเก็บข้อมูลจากรายวิชาที่จัดการเรียนการสอนขึ้นในระดับปริญญาตรี ซึ่งเป็นวิชาในกลุ่มศึกษาทั่วไป จากการจัดการศึกษาพบว่าคำถามสำคัญที่เกิดขึ้นระหว่างการเรียนการสอนในแต่ละภาคการศึกษาคือ การประมาณการผลสัมฤทธิ์ที่จะเกิดขึ้นเมื่อสิ้นสุดภาคการศึกษา ซึ่งคำถามดังกล่าวจะเกิดขึ้นตลอดระยะเวลาการศึกษา จากคำถามดังกล่าวนำมาสู่การแก้ปัญหาด้วยวิธีการพยากรณ์

ผลสัมฤทธิ์ที่จะเกิดขึ้นกับผู้เรียนในอนาคต ผู้วิจัยจึงจัดทำระบบการเรียนออนไลน์ (e-Learning) และนำไปใช้ในการจัดการเรียนการสอนแบบผสมผสาน เพื่อเป็นเครื่องมือสำหรับรวบรวมข้อมูลที่เกิดขึ้นตลอด 6 ภาคการศึกษา (2/2553 ถึง 1/2556) มีจำนวนผู้เรียนทั้งสิ้น 1,605 คน และมีการเก็บรวบรวมข้อมูลปัจจัยต่าง ๆ ของผู้เรียนทุกคนสำหรับใช้ในการวิจัย ชุดข้อมูลที่เก็บรวบรวมมีรายละเอียดดังตารางที่ 1

ตารางที่ 1: รายละเอียดปัจจัยที่ส่งผลต่อผลสัมฤทธิ์

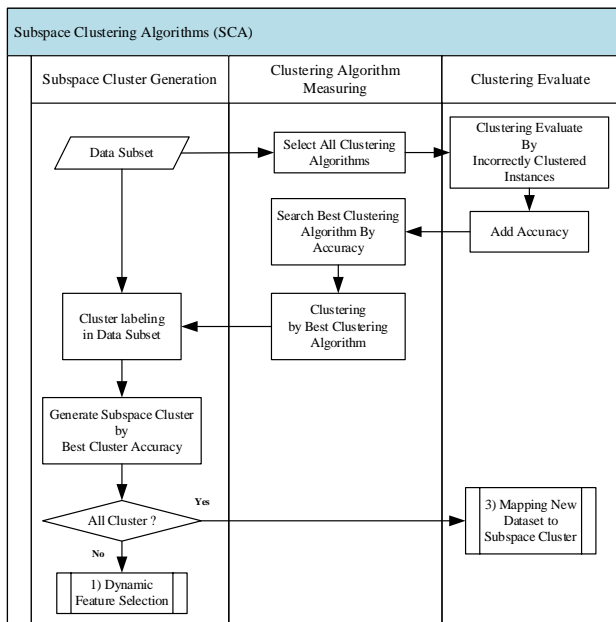
รายการ	ชื่อตัวแปร	หมายเหตุ
1	Major	สาขา
2	Level	ระดับชั้นปี
3	Group.Lec	กลุ่มบรรยาย
4	Group.Lab	กลุ่มปฏิบัติ
5	Semester	ภาคการศึกษา
6	Pre-test	ทดสอบก่อนเรียน
7	FT-1	ทดสอบบทที่ 1
8	FT-2	ทดสอบบทที่ 2
9	FT-3	ทดสอบบทที่ 3
10	FT-4	ทดสอบบทที่ 4
11	FT-5	ทดสอบบทที่ 5
12	FT-6	ทดสอบบทที่ 6
13	FT-7	ทดสอบบทที่ 7
14	FT-8	ทดสอบบทที่ 8
15	Post-test	ทดสอบหลังเรียน
16	Avg.AllTest	ค่าเฉลี่ยของ 1-12
17	Avg.1-8Test	ค่าเฉลี่ยของ 2-11
18	NumOfPass	ค่า 1-12 >= 60%
19	Score.Lab	คะแนนเต็ม 25
20	Score.eLearning	คะแนนเต็ม 15
21	Score.Mit	คะแนนเต็ม 30
22	Score.Final	คะแนนเต็ม 30
23	Grade	8 ระดับ (A,B,...F)

3.3 การสร้างแบบจำลองสำหรับการเลือกปัจจัยที่ส่งผลต่อผลสัมฤทธิ์ทางการเรียน

เมื่อศึกษาในรายละเอียดของค่าคะแนนที่เกิดขึ้นจากกิจกรรมตลอดภาคการศึกษาเทียบกับผลสัมฤทธิ์ทางการศึกษาที่

เกิดขึ้นเมื่อผู้เรียนจบการเรียนรู้ในรายวิชานั้น ๆ จะพบว่า ค่าคะแนนที่เกิดขึ้นจากกิจกรรมตลอดภาคการศึกษา มีความสัมพันธ์โดยตรงต่อผลสัมฤทธิ์ทางการเรียนของผู้เรียน และมีระดับความสัมพันธ์ที่แตกต่างกันในแต่ละประเภทของกิจกรรม อีกทั้งข้อมูลค่าคะแนนของกิจกรรมต่าง ๆ ยังมีลักษณะเป็น ชุดข้อมูลทับซ้อน (Overlap Data Sets) คือ ผู้เรียนที่มีระดับผลสัมฤทธิ์ที่แตกต่างกัน เช่น ผู้เรียนที่ได้เกรด A กับผู้เรียนที่ได้เกรด C มีการทำคะแนนในบางกิจกรรมใกล้เคียงกัน และบางกิจกรรมแตกต่างกัน โดยความแตกต่างอาจมีลักษณะมากกว่าหรือน้อยกว่า การวิจัยเรื่องการจำแนกกลุ่มข้อมูลบนชุดข้อมูลที่มีความทับซ้อน (Overlapping Data Clustering) [8], [9] จะเน้นในการคัดแยกข้อมูลซึ่งมีความเป็นสมาชิกของหลายกลุ่มข้อมูล เพื่อเพิ่มค่าความถูกต้องในการจำแนกข้อมูลให้สูงขึ้น

เมื่อพิจารณาชุดข้อมูลที่ใช้ในงานวิจัยนี้จะพบว่า ลักษณะข้อมูลมีคุณสมบัติสอดคล้องกับอัลกอริทึมการจับกลุ่มบนปริภูมิย่อย (Subspace Clustering Algorithms : SCA) ผู้วิจัยจึงได้นำเสนอกระบวนการแสดงดังภาพที่ 1



ภาพที่ 1: จำแนกกลุ่มบนปริภูมิย่อย (Subspace Clustering Algorithms)

จากภาพที่ 1 แสดงกระบวนการในการสร้างชุดข้อมูลย่อย (Data Subset) ซึ่งจะประกอบไปด้วย Feature ต่าง ๆ ตามโครงสร้างตารางที่ 1 ในขั้นตอนแรกกระบวนการทำงานจะ

เลือกตัวแปรตามลำดับครั้งละ 1 ตัว (1 Feature) และนำตัวแปรดังกล่าวเพิ่มชุดคำตอบ (Target Class) เพื่อส่งต่อไปยังกระบวนการเรียนรู้การจัดกลุ่มข้อมูล (Clustering Algorithm) เมื่อทำการจำแนกกลุ่มข้อมูลเสร็จสิ้น ตัวแปรดังกล่าวจะได้รับการประเมินค่าความถูกต้องของการจำแนกกลุ่ม (Accuracy) เป็นดัชนีชี้วัดความถูกต้อง ซึ่งคำนวณจากการนำกลุ่มที่จำแนกได้จากอัลกอริทึมมาเปรียบเทียบกับค่าคำตอบหรือ Target Class ที่ได้จัดเตรียมไว้ จากนั้นทำการวนซ้ำกระบวนการข้างต้น จนครบทุกตัวแปร และเมื่อครบทุกตัวแปรจะทำการค้นหาตัวแปรที่ให้ค่าความถูกต้องมากที่สุด (Best Feature) เพื่อนำมาสร้างเป็นชุดข้อมูลย่อย (Data Subset) ในขั้นตอนสุดท้ายของกระบวนการ ผู้วิจัยจะนำชุดข้อมูลย่อยที่ได้จากขั้นตอนข้างต้นไปเข้าสู่ อัลกอริทึมการจำแนกกลุ่มบนปริภูมิย่อย (Subspace Clustering Algorithms) สำหรับหาจำนวนตัวแปรที่ส่งผลให้ค่าความถูกต้องของการจำแนกมากที่สุด เพื่อใช้สำหรับเลือกจำนวนตัวแปรในชุดข้อมูลย่อย ที่ส่งผลทำให้ประสิทธิภาพในการจำแนกกลุ่มข้อมูลมากที่สุด

การทดสอบออกเป็น 3 ขั้นตอนคือ 1) นำข้อมูลสำหรับเรียนรู้เข้าสู่ อัลกอริทึมการจำแนกกลุ่มบนปริภูมิย่อย 2) วิเคราะห์หาปัจจัยเพื่อการจัดกลุ่มที่ดีที่สุด แยกตามระดับผลการศึกษา 3) วัดประสิทธิภาพของปริภูมิย่อยด้วยค่าความถูกต้องในการจำแนกกลุ่ม (Correctly Clustered Instances : CCI) โดยการเปรียบเทียบกับ Target Class จำนวน 8 Class (A, B+, ..., F) ของชุดข้อมูลสำหรับเรียนรู้

3.3.1 การเตรียมข้อมูลเข้าสู่ อัลกอริทึม

การจัดเตรียมข้อมูลเพื่อใช้ในการวิเคราะห์อัลกอริทึมเริ่มต้น โดยการปรับค่าเกณฑ์การตัดเกรดให้อยู่ในมาตรฐานเดียวกัน ทั้งนี้เนื่องจากในแต่ละภาคการศึกษามีการตัดเกรดแบบอิงกลุ่มทำให้ค่าเกณฑ์การตัดเกรดมีค่าที่แตกต่างกัน ผลจากการปรับค่าเกณฑ์การตัดเกรด ทำให้จำนวนผู้เรียนในแต่ละเกรดรายละเอียดดังตารางที่ 2

ตารางที่ 2: ข้อมูลจำนวนผู้เรียนและคะแนนเฉลี่ยแบ่งตามเกรด

เกรด	จำนวนผู้เรียน	ร้อยละ	ค่าคะแนนเฉลี่ย
A	69	4%	88.7
B+	257	16%	83.4
B	374	23%	78.0
C+	357	22%	71.6

C	255	16%	64.8
D+	124	8%	58.3
D	75	5%	52.0
F	94	6%	30.2
สรุป	1605		

จากตารางที่ 2 ข้อมูลจะถูกนำไปเข้าสู่ อัลกอริทึมการจัดกลุ่มบนปริภูมิย่อย (Subspace Clustering Algorithms : SCA) หารายจายจากตารางที่ 2 ซึ่งส่งผลดีที่สุดในการจำแนกกลุ่มตามระดับคะแนนเกรด เมื่อนำผลการจำแนกกลุ่มที่ได้ไปเทียบกับจำนวนผู้เรียนในตารางที่ 2

3.3.2 วิเคราะห์หาปัจจัยเพื่อการจัดกลุ่มที่ดีที่สุด

ผลการวิเคราะห์ปัจจัยที่ส่งผลต่อการจัดกลุ่มผ่านอัลกอริทึม SCA มีรายละเอียดดังตารางที่ 3

ตารางที่ 3: ผลการวิเคราะห์ปัจจัยด้วยอัลกอริทึม SCA

ปัจจัย	A	B+	B	C+	C	D+	D	F
Major							✓	
Level			✓	✓		✓	✓	
Group.Lec			✓			✓		
Pre-test		✓			✓		✓	
FT-1							✓	
FT-2							✓	
FT-4		✓				✓		
FT-5						✓	✓	
FT-6	✓						✓	
FT-7		✓						
FT-8						✓		
Post-test						✓	✓	
Avg.AllTest						✓		
Avg.1-8Test						✓		
NumOfPass	✓					✓		✓
Score.Lab						✓	✓	
Score.eLearning		✓			✓	✓	✓	✓
Score.Mit	✓	✓				✓	✓	
Score.Final	✓	✓				✓	✓	
ลำดับการจัดกลุ่ม	2	1	7	8	5	6	3	4

จากตารางที่ 3 แสดงข้อมูลปัจจัยที่ได้จากการวิเคราะห์จากอัลกอริทึม SCA ซึ่งแบ่งผลการทำงานออกได้เป็น 2 ส่วนคือ

จำนวนสมาชิกของตัวแปรในชุดข้อมูล (Feature Set) ที่ทำให้การจัดกลุ่มแต่ละระดับเกรดมีความถูกต้องมากที่สุด และลำดับของการจัดกลุ่มบนปริภูมิย่อยทั้ง 8 กลุ่ม

3.3.3 วัดประสิทธิภาพของปริภูมิย่อย

การพิจารณาถึงประสิทธิภาพที่เกิดขึ้นจากอัลกอริทึม SCA ผู้วิจัยจะนำกลุ่มที่ได้จากการจัดด้วยอัลกอริทึมเทียบกับจำนวนข้อมูลแยกตามกลุ่มตามจำนวนที่แสดงในตารางที่ 2 เพื่อหาค่าความถูกต้อง รายละเอียดดังตารางที่ 4

ตารางที่ 4: ลำดับและค่าความถูกต้องในการจัดกลุ่ม

เกรด	ลำดับการจัดกลุ่ม	จำนวนผู้เรียน (Target)	จำนวนผู้เรียน (SCA)	Accuracy (CCI)
A	2	69	69	100%
B+	1	257	243	95%
B	7	374	235	63%
C+	8	357	349	98%
C	5	255	195	76%
D+	6	124	85	69%
D	3	75	69	92%
F	4	94	86	91%
สรุป		1605	1331	82.92%

จากตารางข้อมูลที่ 3 แสดงลำดับในการจัดกลุ่ม ที่เกิดขึ้นจากอัลกอริทึม SCA โดยระดับคะแนน B+ ถูกจัดกลุ่มได้เป็นอันดับแรกที่ค่าความถูกต้องเท่ากับ 95% ระดับถัดมาได้แก่ ค่าระดับคะแนน A มีค่าความถูกต้องเท่ากับ 100% และ ค่าระดับคะแนน D มีค่าความถูกต้องเท่ากับ 92% ตามลำดับ ทั้งนี้เมื่อพิจารณาประสิทธิภาพโดยรวมในด้านความถูกต้องของการจัดกลุ่ม พบว่า จำนวนข้อมูลทั้งหมดในชุดข้อมูลเท่ากับ 1605 คน อัลกอริทึม SCA มีประสิทธิภาพในการจัดกลุ่มได้ถูกต้อง 1331 คน แยกออกเป็น 8 ระดับคะแนน คิดเป็นร้อยละ 82.92%

4. สรุปผลวิจัยและข้อเสนอแนะ

ผลการวิจัยพบว่าปัจจัยที่สามารถจำแนกกลุ่มข้อมูลได้อย่างมีประสิทธิภาพ และถูกนำมาใช้ในการแยกผู้เรียนตามระดับคะแนนมากที่สุดได้แก่ ค่าคะแนนด้วยการะบวนการเรียนออนไลน์ (Score.eLearning) เป็นปัจจัยของระดับ

คะแนน 5 เกรด ลำดับถัดมาได้แก่ ชั้นปีของผู้เรียน (Level) ค่าคะแนนสอบกลางภาค (Score.Mit) ค่าคะแนนสอบปลายภาค (Score.Final) เป็นปัจจัยของระดับคะแนน 4 เกรด

เมื่อพิจารณาจากค่าประสิทธิภาพของปริภูมิย่อยด้วยค่าความถูกต้องในการจำแนกกลุ่ม (Correctly Clustered Instances : CCI) พบว่า ระดับคะแนน A มีค่าประสิทธิภาพสูงที่สุดเท่ากับ 100% ในการจำแนกกลุ่มผู้เรียน โดยใช้ปัจจัย 4 ปัจจัยได้แก่ จำนวนครั้งในการสอบผ่านแบบทดสอบระหว่างเรียน (NumOfPass) แบบทดสอบประจำบทเรียนที่ 6 (FT-6) คะแนนสอบกลางภาค (Score.Mit) ค่าคะแนนสอบปลายภาค (Score.Final) โดยมีลำดับในการจัดกลุ่มเป็นลำดับที่ 2

ผลการวิจัยที่ได้สามารถนำไปปรับปรุงวิธีการจัดการเรียนการสอนโดยการเพิ่มกิจกรรมที่ส่งผลต่อผู้เรียนในแต่ละระดับเกรด โดยมุ่งเน้นที่การจัดกิจกรรมการเรียนการสอนเสริมให้ตรงกับกลุ่มผู้เรียนมากขึ้น ซึ่งถือเป็นประโยชน์ที่ผู้สอนจะได้รับเพิ่มขึ้น และประโยชน์ที่เกิดกับผู้เรียนจะส่งผลให้ผู้ถูกกระตุ้นด้วยสารสนเทศซึ่งผ่านกระบวนการวิเคราะห์และมีความหนาเชื่อถือ อีกทั้งผู้เรียนยังสามารถเตรียมความพร้อมก่อนการเรียนได้อย่างมีประสิทธิภาพเมื่อได้รับสารสนเทศซึ่งเป็นผลลัพธ์ของงานวิจัยชิ้นนี้

เอกสารอ้างอิง

- [1] C.-F. Tsai, C.-T. Tsai, C.-S. Hung, and P.-S. Hwang, "Data Mining Techniques for Identifying Students at Risk of Failing a Computer Proficiency Test Required for Graduation," *Australasian Journal of Educational Technology*, vol. 27, no. 3, pp. 481–498, 2011.
- [2] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques: Concepts and Techniques*. Elsevier, 2011.
- [3] N. Agarwal, E. Haque, H. Liu, and L. Parsons, "A subspace clustering framework for research group collaboration," *International Journal of Information Technology and Web Engineering*, vol. 1, pp. 35–58, 2006.
- [4] L. Parsons, E. Haque, and H. Liu, "Subspace clustering for high dimensional data: a review," *SIGKDD Explor. Newsl.*, vol. 6, no. 1, pp. 90–105, Jun. 2004.
- [5] J. Macqueen, "Some methods of classification and analysis of multivariate observations," presented at the

Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, 1967, pp. 281–297.

- [6] K. Wagstaff, C. Cardie, S. Rogers, and S. Schrödl, "Constrained K-means Clustering with Background Knowledge," in *Proceedings of the Eighteenth International Conference on Machine Learning*, San Francisco, CA, USA, 2001, pp. 577–584.
- [7] I. H. Witten, E. Frank, and M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques, Third Edition*, 3rd ed. Morgan Kaufmann, 2011.
- [8] H. Lu, Y. Hong, W. N. Street, F. Wang, and H. Tong, "Overlapping Clustering with Sparseness Constraints," in *2012 IEEE 12th International Conference on Data Mining Workshops (ICDMW)*, 2012, pp. 486–494.
- [9] L. O. Mak, G.-W. Ng, G. Lim, and K. Mao, "A merging Fuzzy ART clustering algorithm for overlapping data," in *2011 IEEE Symposium on Foundations of Computational Intelligence (FOCI)*, 2011, pp. 1–6.

การจำแนกน้ำเสียงภาษาไทยเชิงขอร้องและเชิงสั่งการ โดยใช้โครงข่ายประสาทเทียม

Classification of Thai Expressive Speech for Requesting and Ordering using Neural Networks

กานดา รงค์สวัสดิ์ (Kanda Rongsawad)¹ และวัชร ฉัตรวิริยะ (Watchara Chatwiriya)²

ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
s6601120@kmitl.ac.th¹, kcwatcha@kmitl.ac.th²

บทคัดย่อ

จุดมุ่งหมายของงานวิจัยนี้คือการจำแนกลักษณะการ
แสดงออกของเสียงพูดหรือน้ำเสียงเชิงขอร้องและน้ำเสียงเชิง
สั่งการ โดยใช้ลักษณะสำคัญหกประเภทได้แก่ ความยาวของ
การออกเสียง พลังงาน ค่าสูงสุดของเซปสตรัม ค่าสัมประสิทธิ์
เซปสตรัมความถี่แบบเมล ความถี่มูลฐานและความถี่
ฟอร์แมนท์ ประโยชน์โดยตรงของงานวิจัยนี้คือการพัฒนา
ระบบปัญญาประดิษฐ์ด้านการรับรู้การแสดงออกเชิงอารมณ์
ของเสียงและนำลักษณะที่ได้จากงานวิจัยนี้ไปใช้เป็นตัวแปร
สำคัญที่ใช้สังเคราะห์เสียงพูดเชิงขอร้องหรือเชิงสั่งการใน
งานวิจัยเกี่ยวกับการสังเคราะห์เสียงพูดสำหรับการติดต่อกับ
มนุษย์

คำสำคัญ: น้ำเสียงภาษาไทย ลักษณะเสียงพูด การจำแนก
น้ำเสียง โครงข่ายประสาทเทียม

Abstract

The purpose of this paper is classification of Thai
expressive speech for requesting and ordering. Six
groups of selected features are Duration, Energy,
Cepstral Peak Prominence, Mel Frequency Cepstral
Coefficients, Fundamental Frequency and Formant
Frequency. The direct benefits of this research is to
implement an expressive speech classification system. By
using features of this paper as parameter to synthesis
expressive speech which is used in communication for
requesting and ordering.

Keywords: Thai Expressive Speech, Speech Feature,
Classification of Expressive Speech, Neural Networks.

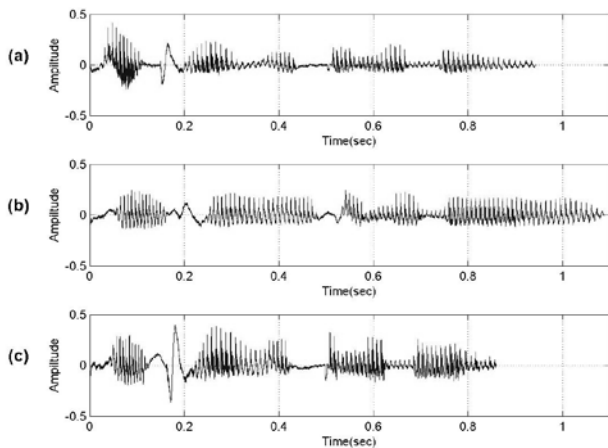
1. บทนำ

กระบวนการสื่อสารด้วยเสียงพูดมีความซับซ้อนมากกว่า
การรู้ความหมายของเนื้อความของประโยคที่พูด เพราะเสียงพูด
ที่เปล่งเนื้อความของประโยคเดียวกัน สามารถสื่อสาร
ความหมายออกมาได้หลายความหมายโดยขึ้นอยู่กับลักษณะ
การออกเสียงที่แตกต่างกัน [1] การเปล่งเสียงที่แตกต่างกัน
ส่งผลต่อปฏิกริยาการตอบสนองของผู้รับฟังที่แตกต่างกัน
ซึ่งกล่าวได้ว่านอกจากเนื้อความแล้ว ต้องเข้าใจการแสดงออก
ของเสียงพูดหรือน้ำเสียงด้วย

เสียงพูดสังเคราะห์ในปัจจุบันสามารถสื่อเนื้อความที่เปล่ง
ออกมาได้อย่างชัดเจน แต่น้ำเสียงยังขาดความเป็นธรรมชาติ
และการแสดงออกเชิงอารมณ์ ดังนั้นปัญหาเกี่ยวกับน้ำเสียงของ
เสียงพูดสังเคราะห์จึงเป็นความท้าทายที่น่าสนใจสำหรับนักวิจัย
ส่วนมาก [2] เนื่องจากการออกเสียงในลักษณะที่แตกต่างกันมี
ความสัมพันธ์ต่อปฏิกริยาการตอบสนองของผู้รับฟังที่แตกต่าง
กันไปด้วย [1]

งานวิจัยนี้มีจุดมุ่งหมายที่ต้องการจำแนกน้ำเสียงภาษาไทยที่
แสดงออกเชิงขอร้องและเชิงสั่งการ โดยการใช้โครงข่าย
ประสาทเทียม และวิเคราะห์หาลักษณะสำคัญที่มีอิทธิพลต่อ
น้ำเสียงเพื่อนำไปประยุกต์ใช้ในงานวิจัยเกี่ยวกับการสังเคราะห์
เสียงพูดเชิงขอร้องหรือสั่งการ งานวิจัยนี้มีสมมุติฐานว่า
น้ำเสียงเชิงขอร้องและน้ำเสียงเชิงสั่งการ มีลักษณะของน้ำเสียง

ที่สำคัญบางตัวเป็นตัวกำหนดน้ำเสียงให้มีลักษณะการเปล่งเสียงเชิงขอร้อง หรือเชิงสั่งการ



ภาพที่ 1: ภาพแสดงสัญญาณเสียงกับเวลา

(a) น้ำเสียงปกติ (b) น้ำเสียงเชิงขอร้อง (c) น้ำเสียงเชิงสั่งการ

จากภาพที่ 1 แสดงสัญญาณเสียงกับเวลาของประโยค “เข้าแถวให้เป็นระเบียบ” ของผู้พูดคนเดียวกัน แต่มีการออกเสียงแตกต่างกันสามประเภท ได้แก่ น้ำเสียงปกติ น้ำเสียงเชิงขอร้องและน้ำเสียงเชิงสั่งการ ดังแสดงในภาพที่ 1(a) 1(b) และ 1(c) ตามลำดับซึ่งเห็นได้ว่าสัญญาณของเสียงทั้งสามประเภทนี้มีลักษณะแตกต่างกัน

บทความนี้แบ่งเป็นห้าส่วนได้แก่ ทบทวนวรรณกรรมที่เกี่ยวข้อง วิเคราะห์ลักษณะที่สนใจ ขั้นตอนการศึกษาวิจัย ผลทดลองและวิเคราะห์ผล และสรุป

2. ทบทวนวรรณกรรมที่เกี่ยวข้อง

ในปัจจุบันมีงานวิจัยที่เกี่ยวข้องกับการวิเคราะห์และสังเคราะห์เสียงพูดที่มีน้ำเสียงหลายประเภทเช่น

ในปี 2005 Eide [3] และคณะนำเสนองานวิจัยเกี่ยวกับการพัฒนาระบบสังเคราะห์เสียงพูดด้วยฐานข้อมูลคำ มีน้ำเสียงแตกต่างกันห้าประเภท คือ น้ำเสียงปกติ, น้ำเสียงเพื่อถ่ายทอดข่าวดี, น้ำเสียงเพื่อถ่ายทอดข่าวร้าย, น้ำเสียงเพื่อถามคำถาม และน้ำเสียงเพื่อเน้นความสำคัญ ในปี 2010 Qin และคณะ [4] นำเสนอการสังเคราะห์เสียงพูดภาษาจีนเพื่อแสดงอารมณ์ด้วยหลักการของพีชชีลอจิกร่วมกับวิธีสังเคราะห์เสียงแบบ HMM ในปี 2012 Alan และคณะ [5] ทดลองสังเคราะห์เสียงพูดภาษาอังกฤษและภาษาเยอรมันด้วยลักษณะด้านสรีรศาสตร์ของเสียงพูดเป็นตัวแปรในการสังเคราะห์เสียงพูด เพื่อปรับปรุง

คุณภาพเสียงพูดให้มีลักษณะการออกเสียงที่เป็นธรรมชาติและสามารถแสดงอารมณ์ได้ ผลการทดลองแสดงให้เห็นว่าลักษณะด้านสรีรศาสตร์ของเสียงพูดเป็นลักษณะที่สามารถนำมาปรับปรุงคุณภาพเสียงให้มีลักษณะการออกเสียงที่เป็นธรรมชาติมากขึ้นได้

งานวิจัยเกี่ยวกับการวิเคราะห์และสังเคราะห์เสียงพูดภาษาไทยเกี่ยวกับน้ำเสียงในปัจจุบัน เป็นงานวิจัยเกี่ยวกับการวิเคราะห์ความแตกต่างของลักษณะการออกเสียงเชิงอารมณ์ เช่น ในปี 2010 และ 2011 Suphattharachai [6][7] นำเสนองานวิจัยเกี่ยวกับการวิเคราะห์เสียงพูดภาษาไทยที่มีน้ำเสียงสี่ประเภท คือ อารมณ์ดีใจ, อารมณ์โกรธ, อ่านหนังสือ และอารมณ์เศร้า งานวิจัยนี้ทำการวิเคราะห์น้ำเสียงด้วยโมเดลคอนทัวร์ความถี่มูลฐานและโมเดลของ Fujisaki ผลลัพธ์จากงานวิจัยพบว่าโมเดลทั้งสองโมเดลสามารถนำมาวิเคราะห์ความแตกต่างของน้ำเสียงทั้งสี่ประเภทได้ แต่ผลลัพธ์ของโมเดลคอนทัวร์ความถี่มูลฐาน มีความผิดพลาดน้อยกว่าโมเดลของ Fujisaki

3. วิเคราะห์ลักษณะที่สนใจ

ลักษณะของเสียงพูดที่งานวิจัยนี้สนใจมีหกประเภทซึ่งเป็นลักษณะที่นิยมใช้ในการวิเคราะห์เสียงพูด แต่ละประเภทมีรายละเอียดดังนี้

1) ความยาวของการออกเสียง เป็นเวลาที่ใช้ในการออกเสียง การเพิ่มความยาวของการออกเสียงมีความสัมพันธ์เกี่ยวข้องกับการเน้นเสียงในประโยค เช่น ภาษาไอซ์แลนด์ เป็นต้น [8]

2) พลังงาน เป็นลักษณะที่มีความสัมพันธ์กับบุคลิกลักษณะและอารมณ์ของผู้พูด และระดับความดังของเสียงแปรผันกับพลังงาน มนุษย์จะตอบสนองกับระดับความดังของเสียงพูดได้ดีกว่าการเปลี่ยนแปลงความถี่ของเสียงพูด แต่ลักษณะนี้ได้รับผลจากเสียงรบกวนจากสิ่งแวดล้อม [9]

3) ค่าสูงสุดของเซปสเตอร์ม หรือ CPP เป็นลักษณะที่มีความสัมพันธ์กับคาบของความถี่มูลฐาน นิยมใช้ลักษณะนี้ในการวิเคราะห์เสียงพูดลมแทรก [9] มีข้อดีคือลักษณะนี้สามารถแสดงระดับคาบ พลังงานของสัญญาณและสามารถแยกเซปสเตอร์มของสัญญาณจากเซปสเตอร์มของสัญญาณรบกวนได้น่าลักษณะนี้ไปประยุกต์ใช้ได้ทั้งสัญญาณที่ไม่ได้ผ่านการกรองและสัญญาณที่ผ่านการกรองแล้ว

4) ค่าสัมประสิทธิ์เซปสตรีมความถี่แบบเมล หรือ MFCC เป็นลักษณะที่นิยมใช้ในงานด้านการวิเคราะห์เสียงพูดอย่างหลากหลาย [9] มีข้อดีคือลักษณะนี้ใช้มาตราวัดแบบเมลซึ่งครอบคลุมย่านความถี่การได้ยินของมนุษย์ ในงานวิจัยนี้ใช้ค่าสัมประสิทธิ์เซปสตรีมความถี่แบบเมลลำดับศูนย์ถึงสิบสอง

5) ความถี่มูลฐาน หรือ F0 คือคอนทัวร์ของพิทช์ที่นิยมนำมาใช้เป็นตัวแทนของระดับสูงต่ำของเสียงพูด ยกตัวอย่างภาษาจีนแมนดารินระดับเสียงสูงต่ำเป็นสิ่งสำคัญต่อความหมายของเสียงที่เปล่งออกมา [9]

6) ความถี่ฟอร์แมนท์ คือความถี่สั้นพ้องที่เกิดจากการสั่นสะเทือนของเส้นเสียงภายในช่องเสียง ฟอร์แมนท์เป็นลักษณะที่สำคัญของเสียงพูดที่มีความเกี่ยวข้องกับการออกเสียงให้ชัดเจน เสียงพูดสามารถมีความถี่ฟอร์แมนท์ได้หลายค่า แต่ที่นิยมนำมาใช้ในการวิเคราะห์ได้แก่ฟอร์แมนท์ลำดับต้น [9] ในงานวิจัยนี้ใช้ความถี่ฟอร์แมนท์ลำดับหนึ่งถึงสี่

ตารางที่ 1: ตารางแสดงกลุ่มลักษณะที่สนใจ

ที่	ประเภทลักษณะ	ลักษณะที่สนใจ
1	ความยาวของการออกเสียง	ความยาวของการออกเสียงแต่ละประโยค
2	พลังงาน	พลังงานของเสียงแต่ละประโยค
3	ค่า สูง สุด ของ เซปสตรีม	ค่าสูงสุดของเซปสตรีมของเสียงแต่ละประโยค
4	ค่าสัมประสิทธิ์เซปสตรีมความถี่แบบเมล	ค่าสัมประสิทธิ์เซปสตรีมความถี่แบบเมลลำดับที่ศูนย์ถึงสิบสองของเสียงแต่ละประโยค
5	ความถี่มูลฐาน	ค่าความถี่มูลฐานของเสียงแต่ละประโยค
6	ความถี่ฟอร์แมนท์	ค่าความถี่ฟอร์แมนท์ลำดับที่หนึ่งถึงสี่ของเสียงแต่ละประโยค

4. ขั้นตอนการศึกษาวิจัย

งานวิจัยนี้มีขั้นตอนการศึกษาวิจัยทั้งหมดสามขั้นตอน ได้แก่ เก็บตัวอย่างข้อมูลเสียงพูด สกัดลักษณะที่สนใจ ทดสอบ

ลักษณะที่ได้ด้วยโครงข่ายประสาทเทียมและสรุปผล มีรายละเอียดดังนี้

4.1 เก็บตัวอย่างข้อมูลเสียงพูด

ขั้นแรกของงานวิจัยทำการเก็บตัวอย่างข้อมูลเสียงพูดที่มีน้ำเสียงที่สนใจสามประเภทคือ น้ำเสียงปกติที่ไม่มีมีการแสดงน้ำเสียง น้ำเสียงเชิงขรื่องที่มีการแสดงเชิงขรื่องหรืออ่อนวอนให้ผู้อื่นทำตามเนื้อความที่เสียงเปล่งออกมา และน้ำเสียงเชิงสั่งการที่มีการแสดงเชิงคำสั่งหรือบังคับให้ผู้อื่นทำตามเนื้อความที่เสียงเปล่งออกมา ในการเก็บตัวอย่างเสียงพูดมีขั้นตอนดังนี้

1) เลือกตัวอย่างผู้พูดที่จะทำการบันทึกเสียง เนื่องจากเพศและช่วงอายุของผู้พูดเป็นปัจจัยที่มีผลที่ทำให้เกิดน้ำเสียงเกิดความแตกต่างกัน ดังนั้นในงานวิจัยนี้เก็บตัวอย่างจากผู้พูดเพียงเพศเดียวคือ เพศชาย ที่มีอายุ 19-25 ปี [11]

2) ออกแบบตัวอย่างประโยคที่ต้องการบันทึกเสียงแสดงตัวอย่างของประโยคดังตารางที่ 2 จากการสังเกตของผู้วิจัย ในขณะที่บันทึกเสียงพบว่า เนื้อความของประโยคที่ให้ผู้พูดเปล่งเสียงมีอิทธิพลต่อลักษณะการเปล่งเสียงของผู้พูด ดังนั้นประโยคที่กำหนดนั้นจึงต้องมีเนื้อความที่สามารถเป็นได้ทั้งเชิงขรื่องและเชิงสั่งการ เพื่อให้ผู้พูดสามารถเปล่งเสียงเพื่อแสดงลักษณะเสียงพูดเชิงขรื่องและสั่งการ ได้

ตารางที่ 2: ตารางแสดงตัวอย่างประโยคที่ใช้เก็บข้อมูล

ที่	ตัวอย่างประโยคที่ใช้เก็บข้อมูล
1	เข้าแถวให้เป็นระเบียบ
2	นำขยะไปทิ้งที่ถังขยะ
3	อย่าส่งเสียงดัง
4	ปิดไฟเมื่อเลิกใช้งาน
5	เดินบนทางเท้าเท่านั้น
6	รอสักครู่
7	อย่าทิ้งขยะในบริเวณนี้
8	อย่านำอาหารเข้ามาในห้อง
9	อยู่ในความสงบ
10	อย่าสูบบุหรี่

3) บันทึกเสียงพูด ให้ผู้พูดแต่ละคนพูดประโยคที่กำหนด จำนวนสามสิบประโยค ในรูปแบบโมโน ค่าแซมปลิงเรต 48 kHz 16 บิต [9]

4) คัดเลือกเสียงพูดที่ได้จากการบันทึก คัดเลือกเสียงพูดที่มีปัญหาทั้ง เช่น เสียงพูดที่มีเสียงรบกวน เสียงพูดที่บันทึกได้ไม่สมบูรณ์ เสียงพูดที่มีน้ำเสียงไม่เด่นชัด ภายหลังจากการคัดเลือกเสียงมีจำนวนตัวอย่างข้อมูลทั้งหมดสี่ร้อยห้าสิบตัวอย่าง

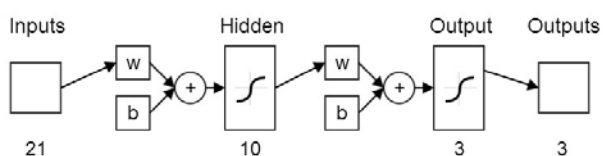
4.2 สกัต์ลักษณะที่สนใจ

ในขั้นตอนนี้ น้ำเสียงพูดที่บันทึกได้จากขั้นตอนที่ 4.1 มาสัต์ลักษณะต่างๆ ที่สนใจตามหัวข้อที่ 3 โดยนำเสียงพูดทั้งประโยคมาแบ่งเสียงออกเป็น ส่วนย่อยขนาดเท่ากัน และทำการ สกัต์ลักษณะที่สนใจด้วยชุดเครื่องมือ VoiceSauce [12] และ VOICEBOX [13] ซึ่งเป็นชุดเครื่องมือที่ใช้สำหรับการ วิเคราะห์เสียงพูด แล้วนำผลมาหาค่าเฉลี่ยของลักษณะแต่ละ ลักษณะในแต่ละประโยค

4.3 ทดสอบลักษณะที่ได้ด้วยโครงข่ายประสาทเทียม

ในขั้นตอนนี้ทดสอบลักษณะที่ได้จากขั้นตอนที่ 4.2 มาแบ่ง ข้อมูลเพื่อใช้ทดสอบการรู้จำน้ำเสียงเชิงขอร้องและเชิงสั่งการ ออกเป็นสามกลุ่มคือ กลุ่มข้อมูลที่ใช้สำหรับฝึกสอน กลุ่มข้อมูล ที่ใช้สำหรับทดสอบ และกลุ่มข้อมูลใช้สำหรับป้องกันเกิด โอเวอร์ฟิตติงในการสอนด้วยอัตราส่วน (8:1:1) และนำมา ตรวจสอบด้วยวิธี 10-Fold Cross Validation

โครงข่ายประสาทเทียมที่ใช้ในการวิจัยใช้โครงข่ายแบบ ซ่อนจำนวนหนึ่งชั้น (single hidden layer) ใช้เรียนรู้แบบ แบคพร็อพกาเกชัน (back propagation) และใช้ฟังก์ชัน ซิกมอยด์เป็นฟังก์ชันกระตุ้นของแต่ละชั้น แสดงในภาพที่ 2 และออกแบบให้เป้าหมายของเสียงพูดแต่ละตัวอย่างเป็นแบบ ไบนารีสามบิต คือเสียงพูดปกติมีเอาท์พุตเป็น [1 0 0] เสียงพูด ขอร้องมีเอาท์พุตเป็น [0 1 0] และเสียงสั่งการมีเอาท์พุตเป็น [0 0 1]



ภาพที่ 2: แสดงโครงสร้างของโครงข่ายประสาทเทียมที่ใช้ในงานวิจัย

จากภาพที่ 2 โครงสร้างของโครงข่ายประสาทเทียมแบ่ง ออกเป็นสามชั้นคือ ชั้นอินพุต ชั้นซ่อน และชั้นเอาท์พุต แต่ละ ชั้นมีขนาดคือ 21 10 และ 3 ตามลำดับ

5. ผลทดลองและวิเคราะห์ผล

ในหัวข้อนี้แบ่งผลการทดลองออกเป็นสองส่วนคือ ผลการ สกัต์ลักษณะที่สนใจจากน้ำเสียงปกติ น้ำเสียงเชิงขอร้องและ น้ำเสียงเชิงสั่งการ และผลการทดสอบการรู้จำน้ำเสียง

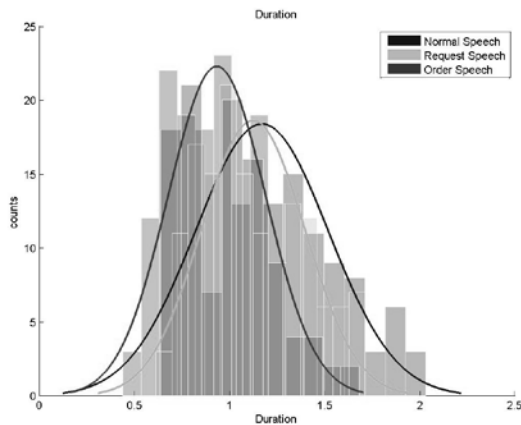
5.1 ผลการสกัต์ลักษณะจากน้ำเสียงปกติ น้ำเสียงเชิงขอร้อง และน้ำเสียงเชิงสั่งการ

จากขั้นตอนที่ 4.2 ทำการสกัต์ลักษณะต่างๆ ที่สนใจ หกประเภทจำนวนยี่สิบเอ็ดลักษณะ แสดงผลการสกัต์ลักษณะ บางส่วนซึ่งมีการกระจายตัวชัดเจน ในตารางที่ 3

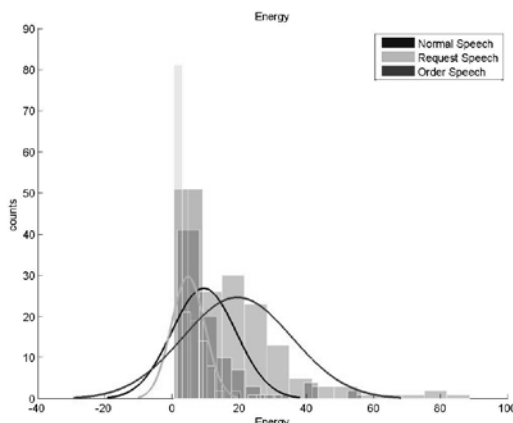
ตารางที่ 3: ตารางแสดงผลสกัต์ที่ได้จากการสกัต์ลักษณะ

ที่	ลักษณะ	ปกติ	ขอร้อง	สั่งการ
		mean	mean	mean
1	ความยาวของการออกเสียง	1.17	1.12	0.93
2	พลังงาน	9.39	4.67	19.38
3	CPP	18.28	17.84	18.86
4	MFCC0	7.64	6.31	9.47
5	F0	129.28	125.41	160.36

จากการวิเคราะห์ผลในตารางที่ 3 เบื้องต้น พบว่าลักษณะ บางตัวของน้ำเสียงแต่ละประเภทมีการกระจายตัวออกจากกัน อย่างชัดเจน แสดงดังฮิสโตแกรมความยาวของการออกเสียง และพลังงาน ตามภาพที่ 3 และภาพที่ 4 ตามลำดับ



ภาพที่ 3: ฮิสโตแกรมการของความยาวของการออกเสียง



ภาพที่ 4: ฮิสโตแกรมของพลังงาน

5.2 ผลการทดสอบการรู้จำน้ำเสียง

จากขั้นตอนที่ 4.3 เมื่อทำการทดสอบการรู้จำน้ำเสียงแต่ละประเภทด้วยโครงข่ายประสาทเทียมและทำการวิเคราะห์ผลการทดลองโดยการสร้างเป็น Confusion Matrix ได้ผลการทดลองแสดงในตารางที่ 4 และ 5

ตารางที่ 4: ตาราง T vb

rainig Confusion Matrix

Output Class	1	73 20.36%	26 7.08%	15 4.11%	64.52%
	2	28 7.89%	79 22.03%	10 2.72%	67.49%
	3	18 5.08%	15 4.22%	95 26.50%	74.01%
		61.08%	66.08%	79.50%	68.89%
		1	2	3	
		Target Class			

ในแนวตั้งแสดงเอาต์พุตที่เกิดขึ้นและในแนวนอนแสดงเอาต์พุตที่ถูกต้อง ในแนวทแยงลงแสดงว่าเอาต์พุตที่เกิดขึ้นที่ตรงกับเอาต์พุตที่ถูกต้อง ตัวเลขเป็นค่าเฉลี่ยที่เกิดขึ้นจากการฝึกสอนโครงข่ายประสาทเทียม ค่าเฉลี่ยความถูกต้องของการฝึกสอนโครงข่ายประสาทเทียม 68.89 เปอร์เซนต์และค่าพรีซิชั่น (precision) และค่ารีคอล (recall) ของการรู้จำน้ำเสียงปกติ 64.52 เปอร์เซนต์ และ 61.08 เปอร์เซนต์ตามลำดับ ค่าพรีซิชั่นและค่ารีคอลของการรู้จำน้ำเสียงเชิงขอร้อง 74.01 เปอร์เซนต์ และ 66.08 เปอร์เซนต์ตามลำดับ และค่าพรีซิชั่นและค่ารีคอลของการรู้จำน้ำเสียงเชิงสั่งการ 74.01 เปอร์เซนต์และ 79.50 เปอร์เซนต์ตามลำดับ

ตารางที่ 5: ตาราง Test Confusion Matrix

Output Class	1	8 18.44%	4 9.56%	2 4.44%	56.85%
	2	4 9.56%	8 18.22%	2 3.78%	57.75%
	3	2 5.33%	3 5.56%	11 25.11%	69.75%
		55.33%	54.67%	75.33%	61.78%
		1	2	3	
		Target Class			

ในแนวตั้งแสดงเอาต์พุตที่เกิดขึ้นและในแนวนอนแสดงเอาต์พุตที่ถูกต้อง ในแนวทแยงลงแสดงว่าเอาต์พุตที่เกิดขึ้นที่ตรงกับเอาต์พุตที่ถูกต้อง ตัวเลขเป็นค่าเฉลี่ยที่เกิดขึ้นจากการทดสอบโครงข่ายประสาทเทียม ค่าเฉลี่ยความถูกต้อง 61.78 เปอร์เซนต์ และค่าพรีซิชั่นและค่ารีคอลของการรู้จำน้ำเสียงปกติ 56.85 เปอร์เซนต์ และ 55.33 เปอร์เซนต์ตามลำดับ ค่าพรีซิชั่นและค่ารีคอลของการรู้จำน้ำเสียงเชิงขอร้อง 57.75 เปอร์เซนต์ และ 54.67 เปอร์เซนต์ตามลำดับและค่าพรีซิชั่นและค่ารีคอลของการรู้จำน้ำเสียงเชิงสั่งการ 69.75 เปอร์เซนต์ และ 75.33 เปอร์เซนต์ตามลำดับ

จากผลการทดลองพบว่าความผิดพลาดของผลลัพธ์ที่มีความผิดพลาดมากที่สุดคือการรู้จำน้ำเสียงปกติและน้ำเสียงขอร้องจากการวิเคราะห์พบว่าความผิดพลาดน่าจะมีสาเหตุมาจากน้ำเสียงทั้งสองประเภทมีความคล้ายคลึงกันมาก ค่าเฉลี่ยความถูกต้องจากการทดสอบโครงข่ายประสาทเทียมของเสียงน้ำเสียงปกติ 56.85 เปอร์เซนต์และน้ำเสียงเชิงขอร้อง 57.75 เปอร์เซนต์ ในขณะที่ค่าเฉลี่ยความถูกต้องจากการทดสอบโครงข่ายประสาทเทียมของเสียงสั่งการ 69.75

เปอร์เซ็นต์ แสดงว่าข้อมูลลักษณะที่ใช้ในงานวิจัยนี้สามารถนำมาใช้เพื่อระบุน้ำเสียงปกติ น้ำเสียงเชิงขร้องและน้ำเสียงเชิงสั่งการได้ถูกต้องอย่างมีนัยสำคัญ

6. สรุป

การศึกษาและจำแนกน้ำเสียงปกติ น้ำเสียงเชิงขร้องและน้ำเสียงเชิงสั่งการของเสียงพูดภาษาไทย มีการออกแบบประโยคที่สามารถใช้น้ำเสียงแบบปกติ ขร้องและสั่งการได้ และเก็บตัวอย่างเสียงจากผู้พูดเพศชาย ได้สกัดลักษณะเสียงจำนวนหกประเภทที่สืบเชื้อลักษณะแล้วนำมาทดสอบความสัมพันธ์ของลักษณะดังกล่าวกับน้ำเสียง โดยการใช้โครงข่ายประสาทเทียม ผลการทดลองแสดงให้เห็นว่าลักษณะเสียงที่ใช้ในงานวิจัยนี้สามารถใช้น้ำเสียงปกติด้วยความถูกต้องเฉลี่ยโดยรวม 56.85 เปอร์เซ็นต์และน้ำเสียงเชิงขร้องด้วยความถูกต้องเฉลี่ยโดยรวม 57.75 เปอร์เซ็นต์ เนื่องมาจากน้ำเสียงปกติและน้ำเสียงเชิงขร้องมีลักษณะคล้ายคลึงกันมาก แต่สามารถระบุน้ำเสียงเชิงสั่งการด้วยความถูกต้องเฉลี่ย 69.75 เปอร์เซ็นต์ อย่างไรก็ตาม เนื่องจากข้อมูลที่ใช้ยังมีจำนวนน้อยจึงน่าจะเพิ่มตัวอย่างในการศึกษาให้มากกว่านี้เพื่อศึกษาว่าจะได้ผลลัพธ์ที่มีแนวโน้มแบบเดียวกันหรือไม่จึงจะสามารถนำผลการการศึกษาไปใช้ประโยชน์ได้ในวงกว้าง

เอกสารอ้างอิง

- [1] Laila Dybkjær and Kristiina Jokinen, *Evaluation of Text*, Springer: Dordrecht, 2007.
- [2] Fang Chen, Holmer Hensen, and Wolfgang Minker, *Speech Technology*, Springer: Dordrecht, 2010.
- [3] E. Eide and et al., "A Corpus-Based Approach to Expressive Speech Synthesis," *5th ISCA Speech Synthesis Workshop*, Pittsburgh, PA, USA, Jun 14- 16, 2004.
- [4] Y. Qin and X. Zhang, "A HMM-Based Fuzzy Affective Model For Emotional Speech Synthesis," *2nd International Conference on Signal Processing Systems (ICSPS)*, Dalian, China, Jul 5- 7, 2010.
- [5] B. W. Alan, and et al., "Articulatory Features for Expressive Speech Synthesis," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Kyoto, Japan, Mar 25- 30, 2012.
- [6] C. Suphattharachai, "Structural Modeling of Fundamental Frequency Contour for Thai Expressive Speech," *Journal of Computer Science*, Vol. 6, No. 3, pp. 330-335, 2010.
- [7] C. Suphattharachai, "Modeling of Fundamental Frequency Contour of Thai Expressive Speech using Fujisaki's Model and Structural Model," *Journal of Computer Science*, Vol. 7, No. 8, pp. 1310-1317, 2011.
- [8] อภิลักษณ์ ธรรมทวีทิกุล, *สัทวิทยาการวิเคราะห์ระบบเสียงใน : ภาษา*, มหาวิทยาลัยเกษตรศาสตร์: กรุงเทพฯ, 2547.
- [9] Ian Mcloughlin, *Applied Speech and Audio Processing: With MATLAB Example*, Cambridge University Press: Cambridge, 2009.
- [10] H. James, C. A. Ronald and E. L. Robert., "Acoustic Correlates of Breathly Vocal Quality," *Journal of Speech and Hearing Research*, Vol. 37, No. 4, pp. 769-778, 1994.
- [11] W. T. Fitch and G. Jay., "Morphology and development of the human vocal tract: A study using magnetic resonance imaging," *The Journal of the Acoustical Society of America*, Vol. 106, No. 3, pp. 1511-1522, 1999.
- [12] Y. Shue, and et al., "VoiceSauce: a program for voice analysis," *International Congress of Phonetic Sciences (ICPhS)*, Hong Kong, China, Aug 17- 21, 2011.
- [13] M. Brookes, "VOICEBOX: Speech Processing Toolbox for MATLAB" [Online]. Available: <http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html>

การตรวจหาคุณลักษณะเด่นด้วยวิธี SIFT สำหรับใบสมุนไพรพื้นบ้าน

A Development of Features Detection using Scale Invariant Feature Transform Approach for Herbal Leaves

รัศม์วริน วรรณศิลป์ (Rakwarinn Wannasin)¹ และกฤตชัย บุญศิวนนท์ (Krittachai Boonsivanon)²

¹ สาขาวิชาแพทย์แผนไทย คณะทรัพยากรธรรมชาติ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี วิทยาเขตสกลนคร

² สาขาวิชาวิศวกรรมคอมพิวเตอร์ คณะอุตสาหกรรมสร้างสรรค์ มหาวิทยาลัยราชภัฏกาฬสินธุ์

¹rakwarinn@outlook.com, ²krittachai@fci.ksu.ac.th

บทคัดย่อ

ปัจจุบันสมุนไพรได้รับความสนใจ และถูกนำมาใช้ในการบำรุงและรักษาสุขภาพร่างกายเพิ่มมากขึ้น แต่สมุนไพรบางชนิดมีลักษณะที่คล้ายคลึงกันก่อให้เกิดความสับสน และไม่สามารถจำแนกสมุนไพรบางชนิดออกจากกันได้ ซึ่งอาจก่อให้เกิดอันตรายต่อสุขภาพร่างกายเมื่อใช้สมุนไพรผิดชนิด และไม่ตรงกับวัตถุประสงค์หรืออาการของโรค งานวิจัยนี้จึงนำวิธีการตรวจหาคุณลักษณะเด่นของใบสมุนไพรพื้นบ้านด้วยวิธี SIFT การวัดประสิทธิภาพของการตรวจหาคุณลักษณะเด่น โดยการทดลองใช้ชุดข้อมูลทดสอบเป็นรูปภาพของใบสมุนไพรทั้งหมด 3 ชนิด ได้แก่ ใบมะรุม ใบสะระแหน่ และใบบัวบก โดยภาพมีขนาด 640x480 พิกเซล จำนวน 3 ชุด ในแต่ละชุดจะมีมุมมอง การหมุนภาพ และความสว่างที่เปลี่ยนแปลง จำนวนชุดละ 30 ภาพ และการใช้วิธีวัดประสิทธิภาพโดยการหาค่า *F-measure* ผลการทดลองพบว่า ค่า *F-measure* ค่าความแม่นยำ (*Precision: P*) และค่าการระลึก (*Recall: R*) ของชุดทดสอบทั้งหมด เท่ากับ 67.0%, 69.0% และ 63.0% ตามลำดับ ส่งผลให้มีประสิทธิภาพในการตรวจหาคุณลักษณะเด่นได้แม่นยำ

คำสำคัญ: วิธี SIFT การตรวจหาคุณลักษณะเด่น สมุนไพรพื้นบ้านไทย การรู้จำ การจับคู่ภาพ

Abstract

Currently, the demand of herbs is increasing rapidly and using various with general medical condition due to their lack of side effects. Many types of herbs have similar features therefore someone cannot classify. The problems following were found a point of confusion and misuse arises from herbs that cause be a lot of harm and adverse herbs reactions to their health. Therefore, in this study is aimed at a development and measure effectiveness of features detection by using SIFT

approach for 3 herbals leaves focus on *Moringa oleifera* Lam, *Metha cordifolia* Opiz and *Centella asiatica* Lin. The three datasets test of resolution are 640 x 480 pixels. Each groups consisted of different viewpoint, rotation and illumination 30 images, which on analysis, using *F-measure*, were identified as *F-measure*, precision and recall rate at 67.0%, 69.0% and 63%, respectively. Our study has demonstrated that it has efficiency can be used in the features detection of herbals leaves accurately.

Keywords: SIFT approach, Features detection, Object-based recognition, Herb, Image matching.

1. บทนำ

ประเทศไทยมีสมุนไพรหลากหลายชนิด บางชนิดนำมาใช้บรรเทารักษาอาการเจ็บป่วยแทนยาแผนปัจจุบัน บางชนิดรับประทานเพื่อบำรุงร่างกายให้แข็งแรง ปัจจุบันสมุนไพรได้รับความสนใจและถูกนำมาใช้เพิ่มมากขึ้น เนื่องจากสมุนไพรเป็นผลิตภัณฑ์ที่ได้จากธรรมชาติและก่อให้เกิดผลข้างเคียงน้อยหรือไม่มีเลย [1] จากที่สมุนไพรมีหลากหลายชนิดบางชนิดมีลักษณะที่คล้ายคลึงกันมากจึงก่อให้เกิดปัญหาคือผู้ใช้สมุนไพรเกิดความสับสนและไม่สามารถจำแนกสมุนไพรบางชนิดออกจากกันได้ ซึ่งอาจก่อให้เกิดอันตรายต่อสุขภาพร่างกายเมื่อใช้สมุนไพรผิดชนิดและไม่ตรงกับความต้องการหรืออาการของโรคที่มีอยู่ นอกจากนี้ปัญหาเกี่ยวกับข้อมูลรูปภาพของสมุนไพรในปัจจุบันที่มีอยู่เป็นจำนวนมากซึ่งบางข้อมูลอาจยังไม่มีมีการจำแนกชนิดที่ชัดเจนหรืออาจเป็นข้อมูลที่ไม่ถูกต้อง จึงยากต่อการจำแนกด้วยตาเปล่าในการเรียนรู้โดยเฉพาะในผู้เรียนที่ขาดความเชี่ยวชาญ ประสบการณ์ และคำแนะนำจากผู้รู้ อาจส่งผลให้การเรียนรู้เกิดความผิดพลาดได้จากปัญหาดังกล่าว ในการจำแนกชนิดของสมุนไพรด้วยตาเปล่ามนุษย์นั้น ผู้ที่จะจำแนกได้ต้องอาศัยความเชี่ยวชาญ

ประสบการณ์และมีความรู้ทางด้านสมุนไพร ซึ่งในบางครั้งยังพบว่าความผิดพลาดในการจำแนกยังคงเกิดขึ้นได้

ดังนั้นการศึกษารุ่นนี้ ผู้วิจัยจึงได้นำเสนอวิธีการตรวจหาคุณลักษณะเด่นด้วยวิธี SIFT สำหรับใบสมุนไพรพื้นบ้าน ซึ่งเป็นวิธีที่ได้รับความนิยมในการนำมาใช้สำหรับรู้จำ (Recognition) เพื่อพัฒนาต้นแบบให้สามารถตรวจหาจับคู่ภาพ และวัดประสิทธิภาพของตรวจหาคุณลักษณะเด่นของรูปภาพของใบสมุนไพรพื้นบ้าน ได้แก่ ใบมะรุม ใบสะระแหน่ และใบบัวบก โดยประโยชน์ที่ได้จากการศึกษารุ่นนี้นั้น สามารถนำไปต่อยอดและประยุกต์ใช้ ในการจำแนกข้อมูลใบสมุนไพรที่มีอยู่เป็นจำนวนมาก (Big data) ในสื่อต่างๆ เช่น อินเทอร์เน็ต หรือตรวจหาคุณลักษณะเด่นในส่วนต่างๆ ของพืชและสมุนไพร รวมทั้งการนำไปใช้ในการเรียนการสอน การเรียนรู้ชนิดของสมุนไพร

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 สมุนไพรพื้นบ้านไทย

การศึกษาเกี่ยวกับลักษณะเด่นโครงสร้างของใบ เพื่อเป็นประโยชน์ในการศึกษาเรียนรู้สำหรับผู้สนใจสมุนไพรในขั้นพื้นฐาน โดยใช้ลักษณะรูปร่างของใบเป็นตัวจำแนก สามารถแบ่งออกเป็น 4 ลักษณะดังนี้

2.1.1 ลักษณะรูปร่างของใบ ได้แก่ รูปกลมแคบ รูปเข็ม รูปแถบ รูปขอบขนาน รูปใบหอก รูปใบหอกกลับ รูปไข่ รูปไข่กลับ รูปสามเหลี่ยม รูปไต รูปวงกลม รูปหัวใจ รูปเลี้ยว และรูปซ้อน

2.1.2 ลักษณะปลายใบ ได้แก่ ปลายใบยาวคล้ายหาง ปลายเป็นติ่งหนาม ปลายแหลม ปลายเรียวแหลม ปลายมน ปลายตัด ปลายเว้ามนุ่ม และปลายเว้าตื้น เป็นต้น

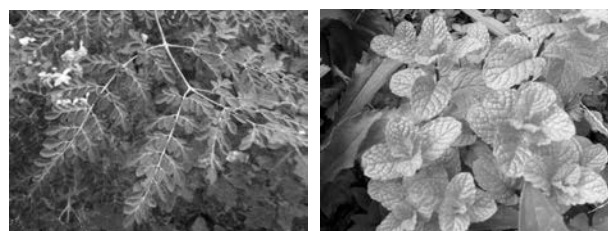
2.1.3 ลักษณะขอบใบ คือ ขอบใบเรียบ ขอบใบเป็นคลื่น ขอบใบหยักมน ขอบใบหยักซี่ฟัน ขอบใบจักฟันเลื่อย ขอบใบเป็นหยัก ขอบใบหยักแบบขนนก ขอบใบหยักเป็นแฉกแบบนิ้วมือ

2.1.4 ลักษณะลักษณะของเส้นใบ คือ เส้นใบแบบขนาน (Parallel veins) พบในพืชใบเลี้ยงเดี่ยว เส้นใบร่างแห (Reticulate veins) พบในพืชใบเลี้ยงคู่

มะรุม (*Moringa oleifera* Lam) เป็นสมุนไพรท้องถิ่นของประเทศทางตะวันตก เขตย่อยของหิมาลัย อินเดีย ปากีสถาน คาบสมุทรเอเชียตะวันตก ทวีปแอฟริกา แหลมอาหรับ [1] จากนั้นแพร่ขยายไปยังประเทศฟิลิปปินส์ กัมพูชา อเมริกากลาง อเมริกาเหนือ อเมริกาใต้ หมู่เกาะแคริบเบียน รวมถึงประเทศไทย ใบมะรุมเป็นสมุนไพรที่อุดมไปด้วยสารต้านอนุมูลอิสระมากที่สุด [2] เป็นยารักษาโรคและให้คุณค่าทางโภชนาการสูง

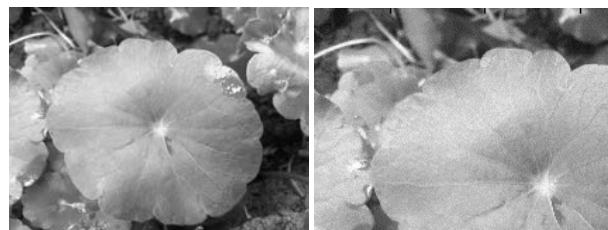
สะระแหน่ (*Metha cordifolia* Opiz) เป็นสมุนไพรท้องถิ่นของชาวทะเลเมดิเตอร์เรเนียนตะวันออก และภูมิภาคเอเชียตะวันตก ใช้เป็นยาบำรุงร่างกาย แก้กระษัย ขับลม ขับเหงื่อ ล้างแผล ยากดประสาทหรือยานอนหลับ เพิ่มประสิทธิภาพของความจำ บรรเทาอาการปวดหัวที่เกิดจากความเครียด [3] ทำให้การรับรู้และอารมณ์ดีขึ้น และเป็นยาคลายความวิตกกังวล

ใบบัวบก (*Centella asiatica* Linn) ปลูกในทวีปเอเชียส่วนใหญ่พบในประเทศอินเดีย ปากีสถาน เกาะมาดากัสการ์ แอฟริกา อเมริกากลาง และโอเชียเนียเขตร้อน ถูกใช้เป็นยาสารพัดประโยชน์มาเป็นเวลา 3,000 กว่าปีในประเทศจีน อินเดีย แอฟริกา ฟิลิปปินส์ ศรีลังกา และเกาะมาดากัสการ์ [4] ในทางการแพทย์พื้นบ้านของชาวเอเชียใบบัวบกเป็นสมุนไพรที่นำมาใช้ในการรักษาโรค และอาการต่างๆ ได้แก่ โรคผิวหนัง ขนาดแผลขนาดเล็ก รอยข่วน แผลไฟไหม้ แผลเป็น แก้อักเสบ โรคเรื้อนกวาง แก้ปวดคลไข้ ขับปัสสาวะ โรคไขข้อ ด้านเชื้อแบคทีเรียและไวรัส โรคมะเร็ง [5] เป็นต้น ดังภาพที่ 1



ก) ใบมะรุม

ข) ใบสะระแหน่



ค) ใบบัวบก

ภาพที่ 1: ใบของสมุนไพร

2.2 การรู้จำด้วยวิธี SIFT

วิธี SIFT (Scale Invariant Feature Transform) [6-7] เป็นวิธีหนึ่งทางด้านทัศนศาสตร์คอมพิวเตอร์ เกี่ยวกับการจับคู่วัตถุจากภาพสองภาพที่มีวัตถุในภาพเหมือนกันที่ได้รับความนิยมนในการศึกษาเพื่อพัฒนาปรับปรุงประสิทธิภาพในการรู้จำหรือจำแนกวัตถุจากภาพ โดยมีขั้นตอนหลักๆ ได้แก่ การตรวจหาคุณลักษณะเด่น (Feature detection) และการจับคู่ภาพ (Key point matching) ที่มีประสิทธิภาพในการรู้จำได้รวดเร็ว โดยทดสอบความทนทานต่อการเปลี่ยนแปลงที่เกิดขึ้นต่างๆ จึงนับได้ว่าเป็นวิธีการพื้นฐานในการพัฒนาปรับปรุงประสิทธิภาพในขั้นตอนการตรวจหาคุณลักษณะเด่นของวิธีอื่นๆ ได้แก่ PCA-SIFT [8-9], GSIFT [9], ASIFT [10] และ SURF [11] เป็นต้น

การคำนวณเพื่อตรวจหาจุดสนใจ (Key point) ในภาพที่มีการเปลี่ยนแปลงเชิง 2 มิติของระดับความเข้มแสงรอบๆจุดสนใจนั้น อัลกอริทึมนี้ได้รับความนิยมโดยมีผู้นำมาศึกษาและพัฒนาปรับปรุงประสิทธิภาพอย่างต่อเนื่อง โดยเฉพาะการพัฒนาในหัวข้อของการรู้จำวัตถุจากภาพ (Object-based recognition) โดยมีขั้นตอนการทำงานหลักๆ 3 ขั้นตอน ได้แก่ การหาและกำหนดจุดสนใจ (Key point detection) การอธิบายจุดสนใจ (Key point description) และการจับคู่ภาพ (Matching) ประกอบด้วย

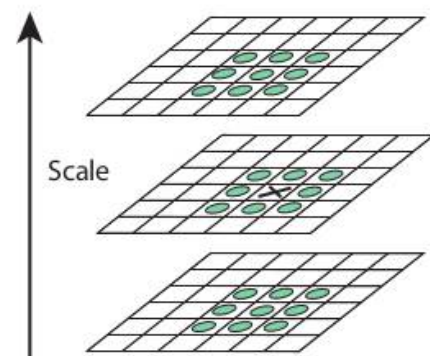
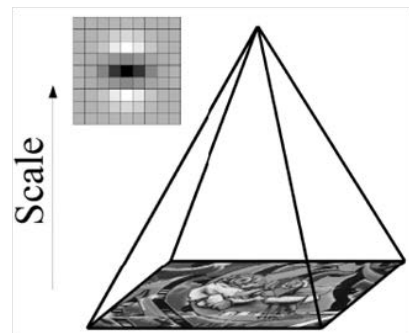
ขั้นตอนที่ 1 Key point detection เป็นขั้นตอนหาปริภูมิในมิติของภาพและระยะทาง (Scale space) โดยการเปลี่ยนแปลงและการเบลอภาพ (Blur) ด้วยฟังก์ชันเกาส์เซียน จากนั้นจึงกำหนดจุดสนใจ (Keypoint) โดยพีรามิด DoG (Difference of Gaussian) ซึ่งถูกปรับปรุงมาจากวิธีเกาส์ลาปลาซ (Gauss-Laplace algorithm) เพื่อลดโอกาสจุดสนใจที่ไม่ชัดเจนหลังจากการเปลี่ยนแปลงสเกล ดังสมการที่ 1 และ 2

$$G(x,y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (1)$$

$$D(x,y, \sigma) = (G(x,y,k\sigma) - G(x,y, \sigma)) * I(x,y) \quad (2)$$

$$= L(x,y,k\sigma) - L(x,y, \sigma)$$

ขั้นตอนที่ 2 Key point localization เป็นขั้นตอนในการหาจุดสนใจ โดยนำค่าที่ได้กำหนดจุดสนใจ 1 จุดภาพ จากนั้นอาศัยการคำนวณค่าของแต่ละจุดภาพรอบๆที่ใกล้เคียงที่สุด (Nearest neighbor) ที่อยู่โดยรอบจุดสนใจ มาทำการคำนวณหาค่าสูงสุด (Local maxima) หรือค่าต่ำสุด (Local minima) ของแต่ละจุดภาพบนปริภูมิ DoG โดยแบ่งเป็น 3 ระดับ จากนั้นจุดสูงสุดหรือจุดต่ำสุดที่ได้จะถูกนำมาจับคู่กับจุดภาพรอบๆอีก 8 จุดภาพใกล้เคียงที่สุดกับจุดสนใจที่ได้กำหนดในข้างต้นในระดับเดียวกัน และทำการคำนวณขนาด 2×9 ส่วน (18 จุดภาพ) ที่อยู่รอบๆ บนปริภูมิ DoG ในสเกลใกล้เคียงในอีก 2 ระดับที่เหลือ ซึ่งวิธีการดังกล่าว คือ การตรวจหาแบบ non-maximal-suppression ขนาด $3 \times 3 \times 3$ ส่วน ดังภาพที่ 2 ซึ่งหากว่าจุดภาพที่ได้เป็นจุดสูงสุด หรือจุดต่ำสุด ก็จะถูกเลือกเป็นจุดสนใจแบบ key point candidate เพื่อใช้ในการคำนวณหาจุดสนใจที่มีแนวโน้มจะเป็นขอบภาพในขั้นตอนถัดไป



ภาพที่ 2: การหาจุดสนใจ (Key point localization) [8]

ขั้นตอนที่ 3 Orientation Assignment เป็นขั้นตอนคำนวณหาขนาด $m(x,y)$ และทิศทาง $\theta(x,y)$ ของจุดที่สนใจเพื่อสร้างรายละเอียดของคุณลักษณะเด่น ดังสมการที่ 3

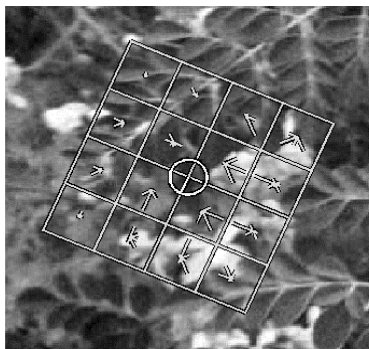
$$m(x,y)=$$

$$\frac{\sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2}}{2} \quad (3)$$

$$\theta(x,y)=\tan^{-1} ((L(x,y+1)-L(x,y-1))/(L(x+1,y)-L(x-1,y)))$$

ขั้นตอนที่ 4 Matching เป็นขั้นตอนสำหรับนำเวกเตอร์อธิบายที่ได้จะถูกนำมาใช้ในขั้นตอนการจับคู่และรู้จำอัตลักษณ์ (Identity) กับภาพต้นฉบับหรือภาพอื่นๆต่อไป โดยใช้วิธีการเปรียบเทียบการจัดของเวกเตอร์ที่แสดงลักษณะเด่นด้วยการวัดแบบยูคลิด (Euclidean distance) กับเวกเตอร์อื่นๆ ดังภาพที่

3



ภาพที่ 3: การตรวจหาจุดสนใจด้วยวิธี SIFT

2.3 งานวิจัยที่เกี่ยวข้อง

การศึกษาและเปรียบเทียบประสิทธิภาพของอัลกอริทึมที่พัฒนาปรับปรุงอัลกอริทึม SIFT [9] ได้แก่ PCA-SIFT, GSIIFT, CSIFT, ASIFT และ SURF และวัดอัตราความถูกต้องในการจับคู่อัลกอริทึมทั้งหมด โดยใช้ชุดภาพทดสอบ (Dataset) ของ Visual Geometry Group และทดสอบความทนทานต่อการเปลี่ยนแปลงที่เกิดขึ้นต่างๆ ได้แก่ ขนาดและการหมุนภาพ การเบลอภาพ การแปลงสีผกผัน และความส่องสว่าง ผลการทดลองพบว่า เวลาเฉลี่ยที่ใช้ในการสกัดคุณลักษณะและจับคู่ภาพ อัลกอริทึม SURF ทำงาน ระดับดีเยี่ยม ในขณะที่การทำงาน ระดับพอใช้ภายใต้การเปลี่ยนแปลงที่เกิดขึ้นอื่นๆ

การศึกษาและเปรียบเทียบอัลกอริทึม SIFT, PCA-SIFT และ SURF [10] โดยใช้วิธีการ K-NN และ RANSAC ในการวิเคราะห์ผลการรู้จำ ในการทดลองความทนทานต่อการได้แก่ ขนาดภาพ การหมุนภาพ การเบลอภาพ และความส่องสว่าง โดยใช้วิธีการวัดค่าประสิทธิภาพ (Repeatability) และเวลาที่ใช้ ผลการทดลองพบว่า SIFT ทำงานได้เร็ว และมีความทนทาน

ต่อการเปลี่ยนแปลงที่เกิดขึ้นต่างๆ เช่น สเกลและการหมุน และความส่องสว่าง เมื่อเทียบกับอัลกอริทึมอื่นๆที่ใช้ในการทดลอง

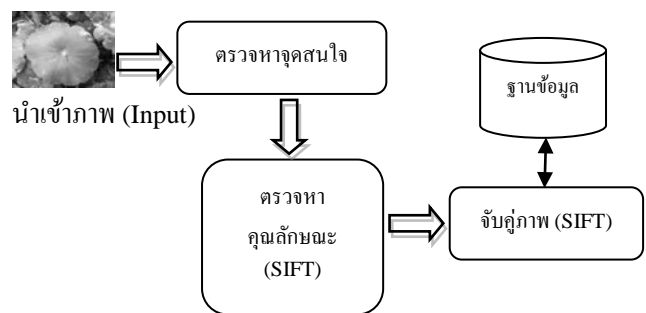
การรู้จำใบสมุนไพรรด้วยเทคนิคการประมวลผลภาพและโครงข่ายประสาทเทียม [12] บนอุปกรณ์ชนิดโปรแกรมฝังตัว โดยใช้คุณลักษณะเด่น คือ รูปทรง และพื้นผิวของใบสมุนไพรร เพื่อพัฒนาระบบจำแนกชนิด และวงศ์พีชสมุนไพรร โดยอัตโนมัติ การทดลองรู้จำสมุนไพรร 20 สายพันธุ์ๆละ 100 ตัวอย่าง ผลการทดลองพบว่า การรู้จำมีความแม่นยำ 98.9%

จากการทบทวนงานวิจัยที่เกี่ยวข้อง พบว่างานวิจัยส่วนใหญ่มีการปรับปรุงประสิทธิภาพในการตรวจหาของวิธี SIFT และใช้การรู้จำใบด้วยเทคนิคการประมวลผลภาพและโครงข่ายประสาทเทียม [12] ยังไม่พบว่าม้งานวิจัยที่ได้นำเอาวิธีการดังกล่าวไปประยุกต์ใช้ในการพัฒนา นอกจากนี้ยังพบว่าอัลกอริทึม SIFT มีประสิทธิภาพในการตรวจหา และจับคู่ภาพได้อย่างรวดเร็ว และทนทานต่อการเปลี่ยนแปลงต่างๆ [9-10] จึงมีความเหมาะสมที่จะนำมาใช้กับงานวิจัยนี้

3. วิธีการดำเนินการวิจัย

3.1 การออกแบบการทดลอง

การพัฒนาการตรวจหาคุณลักษณะเด่นของใบสมุนไพรรที่บ้านด้วยวิธี SIFT ดังภาพที่ 4



ภาพที่ 4: การตรวจหาคุณลักษณะเด่นของใบสมุนไพรรที่บ้านด้วยวิธี SIFT

3.2 การเตรียมการทดลอง

การพัฒนาวิธี SIFT เพื่อใช้สำหรับทดลองการตรวจหาคุณลักษณะเด่น และจับคู่ภาพ อาศัยการเขียนโปรแกรมด้วย Visual C++ และไลบรารี VLFeat [13] โดยทดลองบนเครื่อง

ไมโครคอมพิวเตอร์ (ซีพียู Intel Pentium Dual-core 2.60 GHz, หน่วยความจำ 4 GB) ภายใต้ระบบปฏิบัติการวินโดวส์ 7

3.3 การทดลองการตรวจหาลักษณะเด่น

การทดลองใช้ชุดข้อมูลทดสอบ (Dataset) เป็นรูปภาพของใบสมุนไพรมะนาว 3 ชนิด ได้แก่ ใบมะนาว ใบสะระแหน่ และใบบัวบก ขนาดภาพ 640x480 จุดภาพ จำนวน 3 ชุด ซึ่งในแต่ละชุดจะมีมุมมอง การหมุนภาพ (มุม 90 และ 180 องศา) และความสว่างที่เปลี่ยนแปลง จำนวนชุดละ 30 ภาพ โดยมีการแบ่งกลุ่มภาพเป็น 2 ชุดย่อย ได้แก่ ชุดภาพใช้สอนการรู้จำ จำนวน 60 ภาพ และชุดภาพใช้ทดสอบ จำนวน 30 ภาพ และวิธีวัดประสิทธิภาพโดยการหาค่า F-measure ค่าความแม่นยำ (Precision: P) และค่าการระลึก (Recall: R)

4. ผลการทดลอง

4.1 ผลการทดลองการรู้จำ

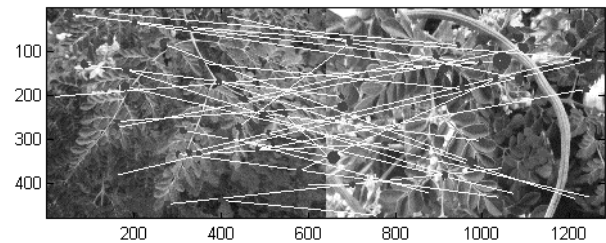
การทดลองและวัดประสิทธิภาพด้วยการหาค่า F-measure ผลการทดลองพบว่า ชุดทดสอบที่สามารถตรวจหาคุณลักษณะได้ดีที่สุด คือ ชุดทดสอบที่ 3 (ใบบัวบก) มีค่า F-measure เท่ากับ 60.0 % และค่าเฉลี่ยของชุดทดสอบทั้งหมด คือ ค่า F-measure ค่าความแม่นยำ (Precision: P) และ ค่าการระลึก (Recall: R) เท่ากับ 67.0%, 69.0% และ 63.0% ตามลำดับ ดังตารางที่ 1

ตารางที่ 1 ค่าเฉลี่ยค่า F-measure ความแม่นยำ และการระลึก

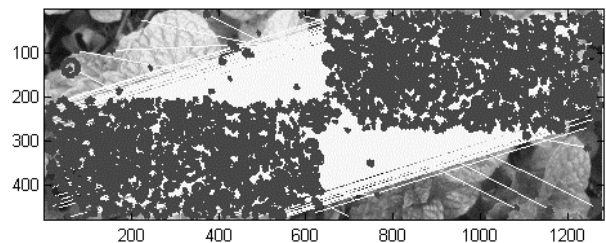
ชุดทดสอบ	F-measure	P	R
ชุดทดสอบที่ 1	0.600	0.604	0.580
ชุดทดสอบที่ 2	0.640	0.675	0.60
ชุดทดสอบที่ 3	0.761	0.790	0.710
ค่าเฉลี่ย	0.67	0.69	0.63

นอกจากนี้จากการทดลองการตรวจหาคุณลักษณะเด่นด้วยการทดสอบความทนทานต่อการเปลี่ยนแปลงที่เกิดขึ้นต่างๆ ได้แก่ มุมมอง การหมุน และความสว่าง พบว่า เมื่อทำการทดสอบโดยใช้ภาพที่ความสว่างของภาพน้อยส่งผลให้การตรวจหาจุดสนใจได้น้อย จึงทำให้อัตราความแม่นยำที่ถูกต้องในการจับคู่ภาพได้

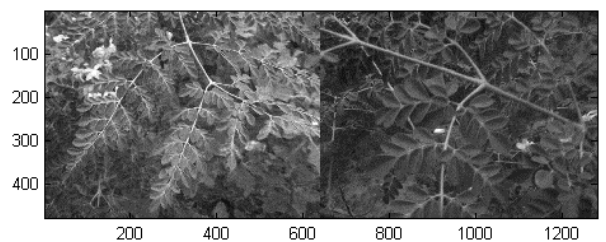
ถูกต้องลดลง ดังนั้นเมื่อนำมาหาค่าเฉลี่ยกับชุดทดสอบทั้งหมด ทำให้ได้ค่า F-measure ค่าด้วยเช่นกัน ในขณะที่การทดสอบความทนทานอื่นๆ สามารถทำงานและมีประสิทธิภาพพอใช้ ดังภาพที่ 5



ก) มุมมอง (ซ้าย:ภาพต้นฉบับ ขวา:ภาพทดสอบ)



ข) การหมุน (ซ้าย:ภาพต้นฉบับ ขวา:ภาพทดสอบ)



ค) ความสว่าง (ซ้าย:ภาพต้นฉบับ ขวา:ภาพทดสอบ)

ภาพที่ 5: ผลการทดสอบด้วยความทนทานต่อการเปลี่ยนแปลงที่เกิดขึ้นต่างๆ

5. สรุป

การพัฒนาการตรวจหาคุณลักษณะเด่นของใบสมุนไพรมะนาวด้วยวิธี SIFT การวัดประสิทธิภาพของตรวจหาคุณลักษณะเด่นของรูปภาพของใบสมุนไพรมะนาว จำนวน 3 ชุด และการหาค่า F-measure ค่าความถูกต้อง และค่าการระลึก ผลการทดลองพบว่า ค่าเฉลี่ยของ ค่า F-measure ความถูกต้อง และค่าการระลึก เท่ากับ 67.0%, 69.0% และ 63.0% ตามลำดับ นอกจากนี้จากการทดลองยังพบว่า เมื่อทำการทดสอบโดยใช้ภาพที่ความสว่างของภาพน้อยส่งผลให้การตรวจหาจุดสนใจได้น้อย จึงทำให้อัตราความแม่นยำที่ถูกต้องในการจับคู่ภาพได้

ถูกต้องลดลง จึงควรพัฒนาปรับปรุงอัลกอริทึมนี้และเลือกใช้
อัลกอริทึมอื่นๆ มาทำการทดลองภายใต้ความทนทานต่อการ
เปลี่ยนแปลงต่างๆ และเปรียบเทียบประสิทธิภาพต่อไป

เอกสารอ้างอิง

- [1] Sh. K. Gediya, R. B. Mistry, U. K. Patel, M. Blessy, and H. N. Jain, "Herbal Plants: Used as a cosmetics," *J. Nat. Prod. Plant Resour.*, vol. 1, no. 1, pp. 24-32, 2011.
- [2] G. Karthivashan, M. T. Fard, P. Arulselvan, F. Abas, and Sh. Fakurazi "Identification of Bioactive Candidate Compounds Responsible for Oxidative Challenge from Hydro-Ethanollic Extract of *Moringa oleifera* Leaves," *Journal of Food Science*, vol. 78, no. 9, pp.1368-1375, 2013.
- [3] R. Sreelekshmi, P. G. Latha, M. M. Arafat, S. Shyamal, V. J. Shine, G. I. Anuja, S. R. Suja, and S. Rajasekharan, "Anti-inflammatory, analgesic and anti-lipid peroxidation studies on stem bark of *Ficus religiosa* Linn," *Natural Product Radianance*, vol. 6, no. 5, pp. 377-381, 2007.
- [4] W. Bylka, P. Z. Awizen, E. S. Sroka, and M. Brzezinska, "Centella asiatica in cosmetology," *Postep Derm Alergol*, vol. 1, pp. 46-49, 2013.
- [5] G. K. Shinomol, K. Muralidhara, and M. M. Bharath, "Exploring the role of Brahmi (*Bacopa monnieri* and *Centella asiatica*) in brain function and therapy," *Recent Pat Endocr Metab Immune Drug Discov*, vol. 5, pp. 33-49, 2011.
- [6] D. G. Lowe, "Distinctive Image Features From Scale-invariant Keypoints," *International Journal of Computer Vision*, Vol. 60, No. 2, pp. 91-110, 2004.
- [7] X. Zhang, Y. Yang, Z. Han, H. Wang and C. Gao, "Object Class Detection: A Survey," *ACM Computing Survey*, Vol. 46, No. 1, pp. 1-54, 2013.
- [8] J. Wu, Z. Cui, V. S. Sheng, P. Zhao, D. Su and S. Gong, "A Comparative Study of SIFT and its Variants," *Measurement Science Review*, Vol. 13, No. 3, pp. 122-131, 2013.
- [9] L. Juan, O. Gwun, "A Comparison of SIFT, PCA-SIFT and SURF," *International Journal of Image Processing*, Vol. 3, No. 4, pp. 143-152, 2009.
- [10] R. Oji, "An Automatic Algorithm for Object Recognition and Detection Based on ASIFT Keypoints" *Signal & Image Processing: An International Journal(SIPIJ)*, Vol. 3, No. 5, pp. 29-40, October 2009.
- [11] H. Bay, T. Tuytelaars and L. V. Gool, "SURF:Speeded-Up Robust Features," *Computer Vision—ECCV 2006: 9th European Conference on Computer Vision, 2006*, 7-13 May 2006, Spinkler, Part II, pp.404-417.
- [12] Z. Husin, A.Y.M. Shakaff, A.H.A. Aziz, R.S.M. Farook, M.N. Jaafar, U. Hashim and A. Harun "Embedded Portable Device for Herb Leaves Recognition using Image Processing Techniques and Neural Network Algorithm," *Computers and Electronics in Agriculture*, Vol. 89, pp. 18-29, 2012.
- [13] A. Vedaldi and B. Fulkerson, "VLFeat-An Open and Portable Library of Computer Vision Algorithms," *Proceedings of the 18th International Conference on Multimedia(MM'10)*, Firenze, Italy, October 25-29, 2010.

**การพยากรณ์อนุกรมเวลาสำหรับปริมาณขยะมูลฝอยของกรุงเทพมหานคร
โดยใช้โครงข่ายประสาทเทียม**
**Time Series Forecasting of the Amount of Solid Waste in Bangkok
using Artificial Neural Network**

ไพบุลย์ ฉัตรทอง (Paiboon Chatthong)¹ มาลีรัตน์ โสคานิล (Maleerat Sodanil)²

และเกียรติศักดิ์ โยชนะนัง (Kiattisak Yochanang)³

^{1,2,3}ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
technocrat_young@yahoo.com¹, msn@kmutnb.ac.th², umph2002tct@hotmail.com³

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อพยากรณ์อนุกรมเวลา สำหรับปริมาณขยะมูลฝอยของกรุงเทพมหานคร โดยใช้โครงข่ายประสาทเทียม ใช้ข้อมูลปริมาณขยะมูลฝอยของกรุงเทพมหานคร ตั้งแต่ ค.ศ. 2545 ถึง ค.ศ. 2556 รวม 130 เดือน ดำเนินการกรองข้อมูลและแปลงค่าข้อมูล เพื่อให้มีคุณสมบัติเป็นอนุกรมเวลาที่คงที่ การสร้างตัวแบบใช้เครื่องมือประเภท Neural Network Time Series ของ MATLAB 8.0 และในการสอนโครงข่ายประสาทเทียมได้ใช้อัลกอริทึมการส่งค่าย้อนกลับวิธีเลเวนเบิร์ก-มาร์ควาร์ดท์พบว่า โครงสร้างที่มีประสิทธิภาพการพยากรณ์ประกอบด้วยจำนวน Neuron ในชั้นซ่อนตัวคือ 25 และจำนวนเวลาข้อนหลังคือ 5 มีค่าความคลาดเคลื่อนการพยากรณ์อนุกรมเวลาอยู่ในระดับต่ำ เมื่อหาค่าพยากรณ์ของปริมาณขยะมูลฝอย ตั้งแต่ปี 2557 - 2561 ปรากฏว่าปริมาณขยะมูลฝอยมีแนวโน้มเพิ่มขึ้น และยังพบว่าโครงข่ายประสาทเทียมมีประสิทธิภาพการพยากรณ์มากกว่า วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล และวิธีการของบ็อกซ์-เจนกินส์

คำสำคัญ: การพยากรณ์อนุกรมเวลา โครงข่ายประสาทเทียม
ขยะมูลฝอย กรุงเทพมหานคร

Abstract

This research aims to forecast time series of the amount of solid waste in Bangkok using artificial neural network. Use of the data amount of solid waste from October 2002 to July 2013 a total of 130 months. Clean the data and convert the data. In order to qualify as a stationary. Building a model with neural network time series tool of Matlab 8.0. Training neural network with the Levenberg-Marquardt algorithm. The results showed that the performance structure consists of neurons number in hidden layer are 25 and delays number are 5. The statistical errors are low level. Find the forecast value of the amount of solid waste since 2014 - 2018. It appears that the amount of solid waste has increased. The neural network is more performance than with Exponential Smoothing and Box-Jenkins.

Keyword: Time Series Forecasting, Artificial Neural Network, Solid Waste, Bangkok

1. บทนำ

กรุงเทพมหานครได้มีการคาดการณ์ปริมาณขยะมูลฝอยโดยใช้อัตราร้อยละที่เพิ่มขึ้นของขยะมูลฝอยร่วมกับ ปริมาณขยะมูลฝอยในปีที่ผ่านมา และใช้อัตราการเกิดขยะมูลฝอยต่อคนต่อ

วันร่วมกับจำนวนประชากรที่คาดคะเน ผลการวิเคราะห์ปรากฏว่าปริมาณขยะมูลฝอยสะสมเพิ่มขึ้นเรื่อย ๆ ตามอนุกรมเวลา ไม่สอดคล้องกับปริมาณขยะมูลฝอยที่เกิดขึ้นจริง

ในต่างประเทศมีการพยากรณ์ปริมาณขยะมูลฝอยชุมชนโดยใช้ตัวแบบความสัมพันธ์ (Causal Model) ทำการพยากรณ์โดยดูความสัมพันธ์ของตัวแปรทางสังคมเศรษฐกิจ [1], [2] ในช่วงเวลาเดียวกัน อย่างไรก็ตามก็จะต้องยอมรับว่าข้อมูลนั้นจะเปลี่ยนแปลงไปตามเวลา มีปฏิสัมพันธ์กับปัจจัยต่าง ๆ การพยากรณ์ปริมาณขยะมูลฝอยด้วยตัวแปรทางสังคมเศรษฐกิจมีความยุ่งยาก จึงควรใช้วิธีการที่แตกต่างเพื่อให้เหมาะสมกับ การเปลี่ยนแปลงของข้อมูลตามกาลเวลา

ดังนั้นการพยากรณ์อนุกรมเวลา จึงเป็นทางเลือกที่ควรดำเนินการ ในงานวิจัยนี้จึงนำอนุกรมเวลาไปใช้ในการพยากรณ์ปริมาณขยะมูลฝอยของกรุงเทพมหานคร โดยใช้โครงข่ายประสาทเทียมเพื่อค้นหารูปแบบและแนวโน้มจากชุดข้อมูล ทั้งนี้จะเปรียบเทียบประสิทธิภาพการพยากรณ์ กับวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล และวิธีการของบ็อกซ์-เจนกินส์โดยใช้ค่าความคลาดเคลื่อนเป็นเกณฑ์

ได้กำหนดสมมติฐานการวิจัยไว้ว่าการพยากรณ์อนุกรมเวลาสำหรับปริมาณขยะมูลฝอยของกรุงเทพมหานคร โดยใช้โครงข่ายประสาทเทียมมีประสิทธิภาพการพยากรณ์มากกว่า วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลและวิธีการของบ็อกซ์-เจนกินส์

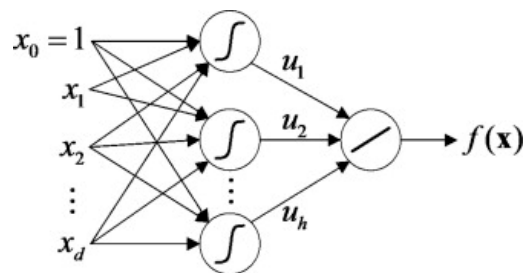
2. วรรณกรรมที่เกี่ยวข้อง

2.1 โครงข่ายประสาทเทียมแบบพอร์เซ็ปตรอนหลายชั้น

เป็นรูปแบบหนึ่งของโครงข่ายประสาทเทียม ที่ใช้สถาปัตยกรรมโครงข่ายประสาทเทียมแบบป้อนไปข้างหน้า ที่มีโครงสร้างเป็นแบบหลาย ๆ ชั้น โดยมีกระบวนการฝึกฝนเป็นแบบมีผู้สอน (Supervise) และใช้ขั้นตอนการส่งค่าย้อนกลับ (Backpropagation) สำหรับการฝึกสอนกระบวนการส่งค่าย้อนกลับนั้นประกอบด้วย 2 ส่วนย่อยคือ การส่งผ่านไปข้างหน้า (Forward Pass) และการส่งผ่านย้อนกลับ (Backward Pass) สำหรับการส่งผ่านไปข้างหน้านั้น ข้อมูลจะผ่านเข้าโครงข่ายประสาทเทียมที่ชั้นข้อมูลเข้า และจะส่งผ่านจากอีกชั้นหนึ่งไปสู่

อีกชั้นหนึ่งจนกระทั่งถึงชั้นข้อมูลออก ส่วนการส่งผ่านย้อนกลับ คำนวณน้ำหนักการเชื่อมต่อจะถูกปรับเปลี่ยนให้สอดคล้อง กับกฎการแก้ข้อผิดพลาด (Error-Correction) คือผลต่างของผลตอบที่แท้จริง (Actual Response) กับผลตอบเป้าหมาย (Target Response) เกิดเป็นสัญญาณผิดพลาด ซึ่งสัญญาณผิดพลาดนี้จะถูกส่งย้อนกลับเข้าสู่โครงข่ายประสาทเทียมในทิศทางตรงกันข้ามกับการเชื่อมต่อ และค่าน้ำหนักของการเชื่อมต่อจะถูกปรับจนกระทั่งผลตอบที่แท้จริงเข้าใกล้ผลตอบเป้าหมาย

หลักการการทำงานของพอร์เซ็ปตรอนหลายชั้นคือในแต่ละชั้นของชั้นซ่อนตัวจะมีฟังก์ชันสำหรับคำนวณ เมื่อได้รับสัญญาณ (Output) จากโหนดในชั้นก่อนหน้านี้ เรียกว่า ฟังก์ชันกระตุ้น (Activation Function) โดยในแต่ละชั้นไม่จำเป็นต้องเป็นฟังก์ชันเดียวกันก็ได้ ชั้นซ่อนตัวนั้นมีหน้าที่สำคัญคือจะพยายามแปลงข้อมูลที่เข้ามาในชั้นนั้น ๆ ให้สามารถแยกแยะความแตกต่างโดยใช้เส้นตรงเส้นเดียว (Linearly Separable) และก่อนที่ข้อมูลจะถูกส่งไปถึงชั้นข้อมูลออก ในบางครั้งอาจจำเป็นต้องใช้ชั้นซ่อนตัวมากกว่า 1 ชั้นในการแปลงข้อมูลให้อยู่ในรูปเส้นตรงเส้นเดียว กระบวนการนี้สามารถอธิบายดังภาพที่ 1 [3]



ภาพที่ 1: การทำงานของโครงข่ายประสาทเทียมแบบพอร์เซ็ปตรอนหลายชั้น

2.2 วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลซ้ำสองครั้ง

วิธีการนี้อาจเรียกว่าวิธีการของโซลท์ ซึ่งจะปรับระดับและแนวโน้มโดยตรงด้วยการใช้ค่าปรับให้เรียบคงที่ (Smoothing Constant) ของแต่ละช่วงเวลาต่างกัน ค่าปรับให้เรียบคงที่เหล่านี้จะประมาณการระดับและแนวโน้มตลอดเวลา เมื่อมีค่าสังเกตใหม่ ดังนั้นจึงมีค่าคงที่ในการทำให้เรียบ 2 ค่า คือ α และ β ตัวแบบของโซลท์มี 3 สมการ ดังนี้ [4]

1) การประมาณระดับ

$$L_t = \alpha Y_t + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad (1)$$

2) การประมาณแนวโน้ม

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \quad (2)$$

3) การพยากรณ์ช่วงเวลา t+p ในอนาคต

$$\hat{Y}_{t+p} = L_t + T_t p \quad (3)$$

โดยที่

L_t	คือ ค่าใหม่หรือค่าประมาณการณปัจจุบัน
α	คือ ค่าปรับให้เรียบคงที่ของการประมาณระดับ
Y_t	คือ ค่าข้อมูลหรือค่าที่แท้จริงของอนุกรมเวลาช่วงเวลา t
T_t	คือ การประมาณแนวโน้ม
β	คือ ค่าปรับให้เรียบคงที่ของการประมาณแนวโน้ม
\hat{Y}_{t+p}	คือ การพยากรณ์ช่วงเวลา t+p ในอนาคต
p	คือ จำนวนของช่วงเวลาที่ต้องการพยากรณ์ในอนาคต

2.3 วิธีการของบ็อกซ์-เจนกินส์

โดยทั่วไปจะเรียกว่าเป็นตัวแทน ARIMA (p, d, q) โดยที่พารามิเตอร์ p, d, q เป็นจำนวนเต็มบวกที่อ้างถึงอัตโนมัติ (AR (p)) บูรณาการ (I) และค่าเฉลี่ยเคลื่อนที่ (MA (q)) สำหรับ AR (p) เป็นรูปแบบที่แสดงว่าค่าสังเกต Y_t ขึ้นอยู่กับค่าของ $Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}$ ส่วน MA (q) เป็นรูปแบบที่แสดงว่าค่าสังเกต Y_t ขึ้นอยู่กับค่าความคลาดเคลื่อน $e_{t-1}, e_{t-2}, \dots, e_{t-q}$ ส่วน Integrated (I) เป็นการหาผลต่างของอนุกรมเวลา เหตุผลที่ต้องหาผลต่างของอนุกรมเวลานี้เนื่องจาก ตัวแบบ ARIMA จะต้องใช้ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่คงที่ (Stationary) ในกรณีข้อมูลอนุกรมเวลาไม่คงที่ จะต้องทำการแปลงข้อมูลอนุกรมเวลาให้มีคุณสมบัติคงที่ โดยการหาผลต่างของข้อมูลอนุกรมเวลา หรือการหาค่าลอการิทึมธรรมชาติของอนุกรมเวลา [4]

2.4 งานวิจัยที่เกี่ยวข้อง

การใช้โครงข่ายประสาทเทียม ในการพยากรณ์ปริมาณขยะมูลฝอยเป็นที่นิยมมากขึ้นในปัจจุบัน ดังงานวิจัยของ Elmira Shamshiry และคณะ [1] มีการใช้จำนวนบุคลากรและค่าใช้จ่ายน้ำมันเชื้อเพลิงเป็นข้อมูลเข้าในตัวแบบ ปรากฏว่าโครงข่ายโครงข่ายแบบ 15-2-1 ได้รับเลือกเป็นโครงข่ายที่ดีที่สุดซึ่งสามารถลดการคำนวณได้ มีบางงานวิจัยได้กำหนดให้ผลิตภัณฑ์มวลรวมภายในประเทศต่อหัว การใช้วัสดุภายในประเทศ และ

ผลิตภาพทรัพยากรเป็นพารามิเตอร์ข้อมูลเข้า ในขณะที่กำหนดให้ การเกิดขยะมูลฝอยเป็นพารามิเตอร์ข้อมูลออก ดังงานวิจัยของ Davor Antanasijević และคณะ [2] ซึ่งมีการใช้สถาปัตยกรรม ทั้งแบบแพร่ย้อนกลับและแบบถดถอย (GRNN) พบว่าแบบ GRNN บรรลุผลดีกว่าในแง่ของค่าสหสัมพันธ์ระหว่างค่าพยากรณ์กับค่าที่วัดได้ในชุดข้อมูลการตรวจสอบ นอกจากนี้ในงานวิจัยของ Kumar, Sudhir J., Subbaiah, Venkata K., and Rao, Prasada P. V. V. [5] ก็ได้กำหนดให้ผลิตภัณฑ์มวลรวมภายในประเทศต่อหัว จำนวนประชากร จำนวนร้อยละของประชากรในเมืองเป็นข้อมูลเข้า โดยให้การเกิดขยะมูลฝอยชุมชนเป็นข้อมูลออก ใช้ฟังก์ชันเรเดียลเบสิส (Radial Basis Function)

ในงานวิจัยของ Ingrida Valavičienė, [6] ได้ใช้ตัวแปรทางสังคมเศรษฐกิจ เช่น GDP สัดส่วนของพลเมืองที่ทำงานในภาคเกษตร และอัตราการตายของเด็ก ฯลฯ เป็นตัวพยากรณ์เพื่อประเมินรูปแบบการผันแปรปริมาณขยะมูลฝอยชุมชน ตามฤดูกาลโดยใช้ตัวแบบ ARIMA และวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลฤดูกาล การพยากรณ์โดยการรวมกันของตัวแบบ ARIMA กับวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลฤดูกาลให้ผลที่ดีที่สุด เมื่อมีแนวโน้มเพิ่มขึ้นของปริมาณขยะมูลฝอยและการเปลี่ยนแปลงตามฤดูกาล ผลของการตรวจสอบชี้ให้เห็นว่าความถูกต้องของตัวแบบผสม ARIMA กับวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลฤดูกาลสูงกว่าของ ARIMA และวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล นอกจากนี้ยังมีงานวิจัยของ I Rimaityte และคณะ [7] ที่ได้สร้างตัวแบบบนฐานตัวแปรทางสังคมเศรษฐกิจ โดยใช้ตัวแบบผสม ARIMA กับวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลฤดูกาลมีข้อสังเกตว่าการพยากรณ์ที่เพิ่มขึ้นทุกปีควรตรวจสอบกับ ข้อมูลที่ได้จากการสร้างตัวแบบการถดถอย

งานวิจัยของ M. Jalili Ghazi Zade, and R. Noori [8] และงานวิจัยของ Himan Shahabi และคณะ [9] ได้ใช้ข้อมูลการเกิดขยะมูลฝอยรายสัปดาห์ $W_t, W_{t-1}, \dots, W_{t-11}$ และจำนวนรถบรรทุกที่บรรทุกขยะมูลฝอย T_t ในการพยากรณ์ปริมาณขยะมูลฝอย W_{t+1} และเพื่อให้บรรลุโครงข่ายที่ดีที่สุดสำหรับการพยากรณ์ขยะมูลฝอย ได้กำหนดโครงข่ายเป็นแบบป้อนไป

ข้างหน้าที่มีสามชั้น และจำนวนของ Neuron ที่แตกต่างกันในชั้นซ่อนตัว ทั้งนี้ใช้ค่า MAE, RMSE และค่า R ในการพิจารณาคัดเลือกตัวแบบที่เหมาะสมพบว่า ผลลัพธ์ที่ดีที่สุดจะเป็นการใช้โครงสร้างแบบ 1-16-13

ในการพยากรณ์อนุกรมเวลาสำหรับปริมาณขยะมูลฝอยของกรุงเทพมหานครนั้น ใช้ข้อมูลปริมาณขยะมูลฝอยแบบรายเดือนเป็นตัวแปรในการศึกษา หากใช้ตัวแปรทางสังคมเศรษฐกิจจะทำให้ตัวแบบไม่มีประสิทธิภาพ เพราะขาดแคลนข้อมูลแบบรายเดือน ทั้งนี้ ข้อมูลเข้าในการพยากรณ์ควรมีจำนวนที่มากพอ

3. วิธีการดำเนินงานวิจัย

3.1 การรวบรวมข้อมูล

ใช้ข้อมูลปริมาณขยะมูลฝอยของกรุงเทพมหานคร ตั้งแต่เดือน ต.ค. 2545 ถึง ก.ค. 2556 รวม 130 เดือนที่รวบรวมจากสำนักสิ่งแวดล้อม กรุงเทพมหานคร

3.2 การกรองข้อมูล (Data Cleaning)

ข้อมูลปริมาณขยะมูลฝอยได้มานั้น ไม่มีค่าของคุณลักษณะขาดหาย ดังนั้นการกรองข้อมูลจะพิจารณาในประเด็นค่าผิดปกติ (Outlier) ที่มีค่าแยกออกจากกลุ่มหรือผิดแผกแตกต่างไปจากข้อมูลค่าอื่น ๆ โดยใช้ค่าคะแนนมาตรฐาน (Z-score) หากข้อมูลค่าใดมีค่าคะแนนมาตรฐานเกิน 2.50 แสดงว่าเป็นค่าผิดปกติในการแก้ไขค่าผิดปกตินั้นจะใช้ค่ามัธยฐาน (เป็นตัวแทนที่ดีกว่าค่าเฉลี่ยในกรณีที่มีค่าผิดปกติ) ของข้อมูลในช่วงเวลา 12 เดือนที่ค่าผิดปกติปรากฏอยู่ไปแทนที่ข้อมูลเดิม

ดำเนินการตรวจสอบค่าผิดปกติ จากค่าคะแนนมาตรฐานของข้อมูลปริมาณขยะมูลฝอยของกรุงเทพมหานคร พบว่าค่าคะแนนมาตรฐานในเดือน ธ.ค. 2554 และ มิ.ย. 2556 มีค่าเกิน 2.50 จึงได้แก้ไขค่าผิดปกติ

3.3 การแปลงค่าข้อมูล

ดำเนินการแปลงค่าข้อมูลปริมาณขยะมูลฝอย ที่แก้ไขค่าผิดปกติแล้วให้มีค่าความแปรปรวนคงที่ก่อน โดยใช้ฟังก์ชันลอการิทึมธรรมชาติ แล้วจึงแปลงให้ค่าเฉลี่ยคงที่โดยใช้ผลต่างอันดับที่ 1 เพื่อให้มีคุณสมบัติเป็นอนุกรมเวลาที่คงที่ ทั้งนี้ข้อมูลปริมาณขยะมูลฝอย ที่แปลงให้ค่าเฉลี่ยคงที่ด้วยผลต่าง

อันดับที่ 1 ของบางช่วงเวลาอาจมีค่าติดลบหากต่อการสื่อความหมายในเชิงเปรียบเทียบ จึงดำเนินการแปลงข้อมูลให้อยู่ในรูปค่าคะแนนที่มีค่าอยู่ระหว่าง 0 - 100

3.4 การสร้างตัวแบบ

3.4.1 แบ่งข้อมูลปริมาณขยะมูลฝอยที่แก้ไขค่าผิดปกติแล้วในรูปค่าคะแนนที่แบบสุ่มออกเป็น 2 ส่วน คือ ส่วนที่ 1 ชุดข้อมูลที่ใช้ในการเรียนรู้ของตัวแบบจำนวน 103 ตัวอย่าง คิดเป็นร้อยละ 80 และส่วนที่ 2 ชุดข้อมูลที่ใช้ในการทดสอบตัวแบบจำนวน 26 ตัวอย่าง คิดเป็นร้อยละ 20

3.4.2 การสร้างตัวแบบโดยใช้โครงข่ายประสาทเทียมใช้โปรแกรม MATLAB 8.0 (R2012b) ทำการเลือกเครื่องมือประเภท Neural Network Time Series และเลือกใช้เทคนิคแก้ปัญหาแบบ Nonlinear Autoregressive (NAR) กำหนดปัญหาในรูปแบบสมการ ดังสมการที่ (4)

$$Y_{(t)} = f(Y_{(t-1)}, \dots, Y_{(t-d)}) \quad (4)$$

การสร้างตัวแบบโดยใช้เทคนิคแก้ปัญหาแบบ NAR มีขั้นตอนดังนี้

ขั้นตอนที่ 1 เลือกข้อมูลที่ใช้ในการเรียนรู้ของตัวแบบ

ขั้นตอนที่ 2 กำหนดจำนวน Neuron ในชั้นซ่อนตัวและจำนวนเวลาซ่อนหลัง (Delay) ทั้งนี้โครงข่ายประสาทเทียมในชั้นซ่อนตัวใช้ฟังก์ชันซิกมอยด์ และชั้นข้อมูลออกใช้ฟังก์ชันเชิงเส้น

ขั้นตอนที่ 3 สอนโครงข่ายประสาทเทียมโดยใช้อัลกอริทึมการส่งค่าย้อนกลับวิธีเลเวนเบิร์ก-มาร์ควาร์ด (Levenberg-Marquardt) ข้อมูลถูกแบ่งออกเป็น 3 ส่วนคือ ส่วนที่ 1 ชุดข้อมูลที่ใช้ในการเรียนรู้ของโครงข่ายประสาทเทียม จำนวนร้อยละ 70 ส่วนที่ 2 ชุดข้อมูลที่ใช้ในการตรวจประเมินจำนวนร้อยละ 15 และส่วนที่ 3 ชุดข้อมูลที่ใช้ในการทดสอบโครงข่ายประสาทเทียมจำนวนร้อยละ 15

ขั้นตอนที่ 4 พิจารณาค่า R หากมีค่าน้อยกว่า 0.75 กลับไปดำเนินการในขั้นตอนที่ 3 ซ้ำอีก (Retrain) และถ้ายังได้โครงข่ายที่ไม่มีประสิทธิภาพการพยากรณ์ก็กลับไปดำเนินการในขั้นตอนที่ 2 ใหม่

ขั้นตอนที่ 5 ทดสอบโครงข่ายประสาทเทียมจากข้อมูลที่ใช้ในการทดสอบตัวแบบ

3.4.3 วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล และวิธีการของบ็อกซ์-เจนกินส์ใช้โปรแกรม SPSS 13.0

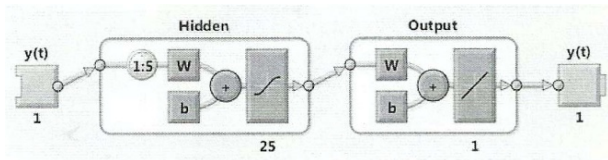
3.5 การประเมินประสิทธิภาพของตัวแบบ

ประเมินตัวแบบในแง่ของประสิทธิภาพการพยากรณ์โดยใช้ค่า RMSE, MAE และ MAPE

4. ผลการดำเนินงานวิจัย

4.1 การพยากรณ์อนุกรมเวลาโดยใช้โครงข่ายประสาทเทียม

โครงสร้างที่มีประสิทธิภาพการพยากรณ์ ประกอบด้วยจำนวน Neuron ในชั้นซ่อนตัว คือ 25 และจำนวนเวลาซ้อนหลัง (Delay) คือ 5 ดังภาพที่ 2



ภาพที่ 2: แผนผังการสร้างตัวแบบโครงข่ายประสาทเทียม

ผลการสร้างตัวแบบได้สมการถดถอย คือ $Output = 0.78 \times Target + 11$ และมีค่า R เท่ากับ 0.8412

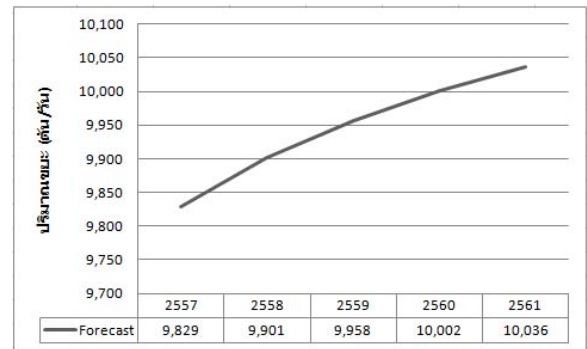
หาค่า RMSE, MAE และ MAPE ได้ผลดังตารางที่ 1

ตารางที่ 1: ค่าความคลาดเคลื่อนการพยากรณ์อนุกรมเวลาโดยใช้โครงข่ายประสาทเทียม

RMSE		MAE		MAPE	
Training	Testing	Training	Testing	Training	Testing
0.0893	0.0812	0.0645	0.0613	0.1286	0.1221

จากตารางที่ 1 จะเห็นได้ว่าค่าความคลาดเคลื่อนการพยากรณ์อนุกรมเวลาทั้งของการสอนตัวแบบและการทดสอบตัวแบบอยู่ในระดับต่ำ หาค่าพยากรณ์ของปริมาณขยะมูลฝอยของกรุงเทพมหานครตั้งแต่ปี 2557 - 2561 ได้ค่าพยากรณ์ดัง

ภาพที่ 3 ทั้งนี้ ได้คำนวณค่าย้อนกลับไปยังค่าตั้งต้นก่อนแปลงค่าข้อมูลเพื่อให้ได้หน่วยเป็นตันต่อวัน



ภาพที่ 3: ค่าพยากรณ์ของปริมาณขยะมูลฝอยของกรุงเทพมหานคร

จากภาพที่ 3 จะเห็นได้ว่าปริมาณขยะมูลฝอยของกรุงเทพมหานครมีแนวโน้มเพิ่มขึ้น

4.2 การพยากรณ์อนุกรมเวลา โดยใช้วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลซ้ำสองครั้งของโฮลท์

จากการวิเคราะห์ข้อมูลโดยเปลี่ยนค่า α และ β ทีละ 0.01 ไปจนถึง 1 พบว่า เมื่อ $L_0 = 49.82208$ และ $T_0 = 0.00565$ ได้ค่า α และ β ที่ทำให้ SSE ต่ำสุด คือ 0.02 และ 0.11 ตามลำดับ

จากค่าเริ่มต้น $L_0 = 49.82208$ และ $T_0 = 0.00565$ หาค่าพยากรณ์ล่วงหน้าทีละหนึ่งหน่วยเวลาโดยใช้สมการที่ (1) สมการที่ (2) และสมการที่ (3)

หาค่า RMSE, MAE และ MAPE ได้ผลดังตารางที่ 2

ตารางที่ 2: ค่าความคลาดเคลื่อนการพยากรณ์อนุกรมเวลาโดยใช้วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลซ้ำสองครั้ง

RMSE		MAE		MAPE	
Training	Testing	Training	Testing	Training	Testing
0.4141	0.3821	0.3134	0.2809	0.6254	0.5592

4.3 การพยากรณ์อนุกรมเวลา โดยใช้วิธีการของบ็อกซ์-เจนกินส์

จากผลการประมาณค่าพารามิเตอร์ในตัวแบบ ARIMA (2,1,0) ที่ไม่มีค่าคงที่ ได้ตัวแบบที่ใช้พยากรณ์ดังสมการที่ (5)

$$\hat{Y}_t = Y_{t-1} - 0.824(Y_{t-1} - Y_{t-2}) - 0.389(Y_{t-2} - Y_{t-3}) \quad (5)$$

หาค่า RMSE, MAE และ MAPE ได้ผลดังตารางที่ 3

ตารางที่ 3: ค่าความคลาดเคลื่อนการพยากรณ์อนุกรมเวลา โดยใช้วิธีการของบ็อกซ์-เจนกินส์

RMSE		MAE		MAPE	
Training	Testing	Training	Testing	Training	Testing
0.4678	0.4401	0.3576	0.3673	0.7131	0.7323

จากตารางที่ 1, 2 และ 3 จะเห็นได้ว่าโครงข่ายประสาทเทียมมีประสิทธิภาพการพยากรณ์มากกว่า วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลและวิธีการของบ็อกซ์-เจนกินส์ เพราะมีค่าความคลาดเคลื่อนอยู่ในระดับต่ำกว่า

5. สรุป

โครงข่ายประสาทเทียม ที่มีประสิทธิภาพการพยากรณ์อนุกรมเวลาสำหรับปริมาณขยะมูลฝอยของกรุงเทพมหานคร มีโครงสร้างประกอบด้วยจำนวน Neuron ในชั้นซ่อนตัว คือ 25 และจำนวนเวลาย้อนหลัง คือ 5 ได้สมการถดถอย คือ $Output = 0.78 \times Target + 11$ และเมื่อหาค่า RMSE, MAE และ MAPE พบว่าค่าเหล่านี้อยู่ในระดับต่ำ หาค่าพยากรณ์ของปริมาณขยะมูลฝอยตั้งแต่ปี 2557 - 2561 พบว่าปริมาณขยะมูลฝอยมีแนวโน้มเพิ่มขึ้น เมื่อเปรียบเทียบกับประสิทธิภาพการพยากรณ์พบว่าโครงข่ายประสาทเทียมมีประสิทธิภาพมากกว่า วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล และวิธีการของบ็อกซ์-เจนกินส์ เป็นไปตามสมมติฐานการวิจัย สาเหตุสืบเนื่องจากโครงข่ายประสาทเทียมเป็นระบบการปรับตัว ที่มีการเปลี่ยนแปลงโครงสร้างในช่วงของการเรียนรู้

เอกสารอ้างอิง

- [1] Shamsiry, Elmira, et al. "Forecasting Generation Waste Using Artificial Neural Networks," *Proceedings of the 2011 International Conference on Artificial Intelligence (ICAI 2011)*, Las Vegas, NV, USA, July 18-21, 2011.
- [2] Antanasijevic', Davor, et al. "The Forecasting of Municipal Waste Generation Using Artificial Neural Networks and Sustainability Indicators," *Sustainability Science*. vol. 8, no. 1, pp. 37-46, 2013.
- [3] Wikipedia. *โครงข่ายประสาทเทียม*. [ออนไลน์] 2556. จาก <http://th.wikipedia.org/wiki/โครงข่ายประสาทเทียม> [สืบค้นวันที่ 15 มี.ค. 2556]
- [4] อัครพงศ์ อ้นทอง. *คู่มือการใช้โปรแกรม EViews เบื้องต้น: สำหรับการวิเคราะห์ทางเศรษฐมิติ*. เชียงใหม่ : สถาบันวิจัยสังคม มหาวิทยาลัยเชียงใหม่, 2550.
- [5] Kumar, Sudhir J., Subbaiah, Venkata K., and Rao, Prasada P. V. V. "Prediction of Municipal Solid Waste with RBF NetWork- A Case Study of Eluru, A.P, India." *International Journal of Innovation, Management and Technology*, vol. 2, no. 3, pp. 238-243, 2011.
- [6] Valavičienė, Ingrida. *The Impact of Seasonal Variation of Municipal Waste Generation on Waste Management System Indicators*. Doctoral Dissertation, Faculty of Chemical Technology, Kaunas University of Technology, 2012.
- [7] Rimaityte, I, et al. "Application and Evaluation of Forecasting Methods for Municipal Solid Waste Generation in An Eastern-European City." *Waste Management & Research*, vol. 30, no. 1, pp. 89-98, 2012.
- [8] Jalili Ghazi Zade, M. and Noori, R. "Prediction of Municipal Solid Waste Generation by Use of Artificial Neural Network: A Case Study of Mashhad." *International Journal of Environmental Research*, vol. 2, no. 1, pp. 13-22, 2008.
- [9] Shahabi, Himan, et al. "Application of Artificial Neural Network in Prediction of Municipal Solid Waste Generation (Case Study: Saqqez City in Kurdistan Province)." *World Applied Sciences Journal*, vol. 20, no. 2, pp. 336-343, 2012.

การพัฒนาเกม White Horse เพื่อการเรียนรู้ภาษาอังกฤษขั้นพื้นฐาน

บนระบบปฏิบัติการแอนดรอยด์

Development of White Horse Game to Basic English Learning on Android

वलกรณ์ ภาสากานนท์ (Walongkorn Pasakanon)¹ และ จงกล จันทร์เรือง (Jongkol Janruang)²

^{1,2}สาขาวิชาเทคโนโลยีสารสนเทศและการสื่อสาร คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน
walong_aon@hotmail.com¹, jj@sci.rmuti.ac.th²

บทคัดย่อ

เกมเพื่อการเรียนรู้บนระบบปฏิบัติการแอนดรอยด์กำลังได้รับความนิยมอย่างมากในปัจจุบัน ดังนั้นรายงานวิจัยฉบับนี้จึงนำเสนอเกม White Horse เพื่อการเรียนรู้ภาษาอังกฤษขั้นพื้นฐานของนักเรียนชั้นประถมศึกษาปีที่ 1-3 โดยนำแนวคิดการเรียนรู้ภาษาอังกฤษแบบ Phonics มาผสมผสานกับรูปภาพและเทคนิคการสร้างเกมเพื่อการเรียนรู้สำหรับเด็ก เพื่อส่งเสริมทักษะภาษาอังกฤษด้านการฟัง การพูด การอ่าน การเขียน การสะกดคำ และเสริมสร้างความสนุกสนานเพลิดเพลินในการเรียนรู้ภาษาอังกฤษผ่านเกม ก่อให้เกิดการเรียนรู้แบบไม่รู้ตัว พร้อมกับการสร้างความมั่นใจด้านภาษาอังกฤษให้กับนักเรียน

คำสำคัญ: เกมส่งเสริมการเรียนรู้ภาษาอังกฤษ เกม White Horse เกมบนแอนดรอยด์ เกมผจญภัย

Abstract

Learning games on the Android operating system is gaining popularity today. Therefore, this research report is to present White Hoses game for learning basics English of Grade 1-3 students by introducing the concept of learning English Phonics combined with pictures and techniques for creating games. Learning for children to promote English language skills of listening, speaking, reading, writing and spelling, and build a fun class to learn English through games cause learning unconsciously along with ensuring the English language to students.

Keyword: Games for Learning English, White Horse Game, Games on Android, Adventure Game

1. บทนำ

การพัฒนาทรัพยากรมนุษย์ ด้วยการสนับสนุนการศึกษาพร้อมกับสนับสนุนการใช้ประโยชน์จากเทคโนโลยีสารสนเทศเป็นหนึ่งในลักษณะงานที่สำคัญตามแผนงานการจัดตั้งประชาคมสังคมและวัฒนธรรมแห่งประชาคมอาเซียน [1] ดังนั้นประเทศไทยต้องมีการยกระดับความเจริญก้าวหน้าด้านเทคโนโลยีสารสนเทศและการสื่อสารเพื่อการศึกษา ในกรณีนี้เองรัฐบาลในยุคปัจจุบันจึงมีนโยบายแจก Tablet ให้แก่นักเรียนชั้นประถมศึกษาปีที่ 1 ทุกคน เนื่องจากรัฐบาลเล็งเห็นความสำคัญของการใช้เทคโนโลยีสารสนเทศและการสื่อสารมาช่วยพัฒนาศักยภาพการศึกษาของเด็กไทยให้ทัดเทียมกับประเทศต่าง ๆ ในกลุ่มสมาชิกประชาคมอาเซียน โดยสามารถเข้าถึงแหล่งเรียนรู้และองค์ความรู้ต่าง ๆ ที่มีอยู่ ทำให้ผู้เรียนมีโอกาสศึกษาหาความรู้ ฝึกปฏิบัติ และสร้างองค์ความรู้ต่าง ๆ ได้ด้วยตัวเอง [2][3] ก่อให้เกิดความทัดเทียมทั้งด้านสิทธิและสวัสดิการสำหรับกลุ่มด้อยโอกาสและกลุ่มที่อ่อนแอ ดังเป้าหมายและพันธกรณี ASCC Blueprint โดยกระทรวงพัฒนาสังคมและความมั่นคงของมนุษย์ [4]

ภาษาอังกฤษมีความสำคัญควรจัดให้มีการเรียนการสอนที่มีประสิทธิภาพ แต่การเรียนการสอนวิชาภาษาอังกฤษในประเทศไทยก็ยังไม่ได้ผลดีเท่าที่ควร [5] อย่างไรก็ตาม ในปี 2552 อินทรา ศรีประสิทธิ์ พัฒนารูปแบบการเรียนรู้ภาษาอังกฤษ 7 ขั้นตอน [6] ซึ่งเกิดจากการบูรณาการการเรียนการสอนแบบ phonics approach กับ whole language approach เพื่อแก้ปัญหาคนไทยภาษาอังกฤษบกพร่อง เนื่องจากไม่สามารถเชื่อมโยงหน่วยเสียงของบทพูดกับตัวอักษร หรือกลุ่มตัวอักษร ทำให้ฟังไม่รู้เรื่อง ซึ่งรายงานวิจัยของ อินทรา ศรีประสิทธิ์ มี

ความสอดคล้องกับงานของ เบญจพร ปุญญะเพ็ชร และ ชลลดา เลหาวิริยานนท์ [7] ที่เสนอแนะว่านักเรียนที่มีความสามารถทางภาษาอังกฤษต่ำกว่าเกณฑ์นั้น ควรต้องมีการเสริมประสบการณ์หรือฝึกฝนภาษาอังกฤษ ด้วยการสร้างกิจกรรมที่หลากหลาย ร่วมกับการเรียนรู้แบบสอนให้รู้จักคำผ่านการใช้รูปภาพเป็นสื่อ ในหนังสือ หรือบัตรคำมิใช่เพียงการอาศัยทฤษฎีของ phonics approach เพียงอย่างเดียว

จากความสำเร็จเรื่องของการเตรียมความพร้อมสู่ประชาคมอาเซียน และปัญหาการเรียนรู้ภาษาอังกฤษของคนไทย ผู้วิจัยจึงมีแนวคิดในการพัฒนาเกม White Horse เพื่อพัฒนาการเรียนรู้ภาษาอังกฤษขั้นพื้นฐานของนักเรียนไทยในชนบท ระดับชั้นประถมศึกษาปีที่ 1 - 3 ด้วยการผสมผสานทฤษฎีการเรียนรู้ภาษาอังกฤษแบบ 7 ขั้นตอนของ อินทรา ศรีประสิทธิ์ กับสอนให้รู้จักคำผ่านรูปภาพของ เบญจพร ปุญญะเพ็ชร และ ชลลดา เลหาวิริยานนท์ ร่วมกับการออกแบบเกมเพื่อการเรียนรู้ที่ดี จักนำมาซึ่งประสิทธิภาพในการเรียนรู้ พร้อมทั้งช่วยลดข้อจำกัดของบุคลากร ความไม่พร้อมของโรงเรียนหรือนักเรียน โดยสิ่งสำคัญที่สุดของการสร้างเกมเพื่อการเรียนรู้ [8] คือการพยายามสร้างการเรียนรู้ด้วยตนเองให้กับผู้เรียน

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

นายวิสิทธิ์ กาญจนธุม [9] พัฒนาเกมชื่อ Bah-Bah มีลักษณะปกป้องหอคอย แบบผู้เล่นเดี่ยว สำหรับเล่นบน Smartphone แบบสัมผัส โดยใช้โปรแกรม Eclipse IDE for Java Developers (Version Indigo Service Release 2) เป็นเครื่องมือในการพัฒนา และใช้ Paint Tool SAI ในการพัฒนากราฟฟิกส์ในการพัฒนาเกม Bah-Bah ผู้วิจัยเน้นความตื่นตัวด้วยการให้ผู้เล่นเป็นผู้สั่งการทุกอย่าง อีกทั้งยังออกแบบให้ผู้เล่นสามารถเอาตัวรอดได้ระยะหนึ่ง แล้วจึงมีร้านค้าปรากฏขึ้นและผู้เล่นจะสามารถซื้ออุปกรณ์เพิ่มขึ้น เพื่อใช้ในด่านถัดไปโดยใช้เงินในเกม ลักษณะการเล่นแบบนี้ ทำให้ผู้เล่นมีความสนุกสนานเพลิดเพลิน เนื่องจากมีความตื่นตัว เกมเล่นไม่ยากจนเกินไป สามารถเอาตัวรอดได้ง่าย ดังนั้นผู้วิจัยจึงนำผลการทดลองของเกมนี้มาปรับใช้ในการพัฒนาเกมเพื่อการเรียนรู้ภาษาอังกฤษ สำหรับนักเรียนระดับประถมศึกษาปีที่ 1 - 3 รวมถึงการใช้เครื่องมือในการพัฒนาเกมของงานวิจัยดังกล่าวมาประยุกต์ใช้

จากการศึกษาวิจัยด้านการบูรณาการการศึกษา (Education) ร่วมกับความบันเทิง (Entertainment) ก่อให้เกิดเป็นความหมายว่าเป็นการให้ความบันเทิงแต่สอดแทรกสาระการเรียนรู้เข้าไปในความบันเทิงนั้น (Edutainment) ด้วยแนวคิดนี้จึงมีการพัฒนาต้นแบบเกมสำหรับการผลิตซอฟต์แวร์เกมผจญภัย ซึ่งมีหัวข้อศีลธรรมตามหลักสูตรการศึกษาของมาเลเซียประกอบอยู่ด้วยในเกม [10] และงานวิจัยของ [11] พบว่าเด็กมีการพัฒนาการที่ดีและเป็นกระบวนการที่ให้นักเรียนมีส่วนร่วม สร้างปฏิสัมพันธ์แบบง่าย ๆ เป็นธรรมชาติ นักเรียนมีความสุขในการเรียนรู้ โดยงานวิจัยทั้งสองฉบับมีการนำสื่อมัลติมีเดียมาใช้ในการออกแบบ Gerard Veenhof [12] ต้องการพัฒนากेमบนมือถือ เพื่อส่งเสริมให้ผู้เรียนมีส่วนร่วมในการเรียนรู้และสร้างแรงบันดาลใจให้กับผู้เรียน ทั้งนี้ผลการทดลองของเขาพบว่าเกม ZooQuest สามารถนำมาสร้างประโยชน์ต่อผู้เรียน เพื่อการฝึกทักษะนอกห้องเรียน อย่างไรก็ตามเกม ZooQuest ไม่ได้ถูกพัฒนาขึ้นด้วยการ Programming แต่เป็นการใช้ open source Android game Andor's Trail มาดัดแปลงเพื่อให้ MEL application กลายเป็นเกมเพื่อการเรียนรู้

3. กรอบแนวคิด

จากทฤษฎีการเรียนรู้ภาษาอังกฤษเรื่อง Phonics และการสอนภาษาอังกฤษด้วยภาพ/บัตรคำและทฤษฎีการพัฒนาเกม (whole language approach) เพื่อการเรียนรู้บนระบบปฏิบัติการแอนดรอยด์ ผู้วิจัยจึงศึกษาปัญหาการเรียนรู้วิชาภาษาอังกฤษด้วยการสอบถามครูและนักเรียนชั้นประถมศึกษาปีที่ 1 - 3 ของโรงเรียนบ้านสาลา ตำบลคงใหญ่ อำเภอบึงสามพัน จังหวัดนครราชสีมา งานวิจัยฉบับนี้จึงมีวัตถุประสงค์เพื่อ

1. พัฒนาเกม White Horse เพื่อการเรียนรู้ภาษาอังกฤษขั้นพื้นฐานสำหรับนักเรียนชั้นประถมศึกษาปีที่ 1 - 3 บนระบบปฏิบัติการแอนดรอยด์

2. ศึกษาประสิทธิภาพและประสิทธิผลของเกม

การวัดประสิทธิภาพและประสิทธิผลของเกม ผู้วิจัยใช้แบบสอบถาม แบบประเมินและแบบสังเกตในการวัดระดับความพึงพอใจ ของนักเรียน รวมถึงวัดพัฒนาการ ความรู้ ที่เพิ่มขึ้นของนักเรียนชั้นประถมศึกษาปีที่ 1-3 จำนวน 24 คน มาเป็นตัวกำหนดค่าประสิทธิภาพและประสิทธิผลของเกม White Horse

4. การออกแบบเกม White Horse

4.1 ขั้นตอนการวิเคราะห์เนื้อหา

ผู้วิจัยได้รวบรวมข้อมูลจากการสัมภาษณ์อาจารย์ผู้สอน โรงเรียนบ้านศาลา และผู้เชี่ยวชาญด้านภาษาอังกฤษ เพื่อศึกษา ปัญหาการเรียนการสอนวิชาภาษาอังกฤษ ระดับชั้น ประถมศึกษาปีที่ 1-3 และสัมภาษณ์นักเรียน เพื่อศึกษา ระดับ ความรู้และทดสอบทักษะภาษาอังกฤษ ตามแผนการสอนวิชา ภาษาอังกฤษของโรงเรียนบ้านศาลา จากการศึกษาข้อมูล เบื้องต้นและการวิเคราะห์ปัญหา รวมถึงความต้องการของ อาจารย์และนักเรียน โรงเรียนบ้านศาลา ผู้วิจัยนำผลการศึกษาดังกล่าว มาเขียนสรุปผลรายงานดังรายละเอียดต่อไปนี้

1) ระดับประถมศึกษาปีที่ 1 นักเรียนควรมีความรู้ ภาษาอังกฤษตามแผนการสอน ด้วยจำนวนคำศัพท์ 150-200 คำ เช่น สี ตัวเลข สิ่งรอบตัวในชั้นเรียน การถามตอบภาษาอังกฤษ ได้ มีความสนุกสนานจากการเรียน เช่น เล่นเกม ร้องเพลง สามารถรับฟังและเข้าใจสถานการณ์ง่าย ๆ ได้

2) ระดับประถมศึกษาปีที่ 2 นักเรียนควรมีความรู้ ภาษาอังกฤษตามแผนการสอน ด้วยจำนวนคำศัพท์ 250-300 คำ เช่น อุปกรณ์การเรียน สิ่งแวดล้อมรอบตัว ประยุกต์ใช้ ภาษาอังกฤษในชีวิตประจำวันได้ ใช้ประโยคภาษาอังกฤษในการเขียนหรือถามสิ่งที่ต้องการ เข้าใจภาษาอังกฤษที่เกี่ยวกับ วิชาอื่นที่เรียนได้ มีความสนุกสนานเพลิดเพลินจากการเรียน

3) ระดับประถมศึกษาปีที่ 3 นักเรียนควรมีความรู้ ภาษาอังกฤษตามแผนการสอน ด้วยจำนวนคำศัพท์ 350-400 คำ เช่น ชื่อประเทศต่าง ๆ คำศัพท์นอกชั้นเรียน ใช้ประโยค ภาษาอังกฤษในการสื่อสารได้ นำกิจกรรมการเรียนรู้ ภาษาอังกฤษ เช่น การร้องเพลง การเล่นเกม มาศึกษาหาความรู้เพิ่มเติมได้ มีความสนุกสนานเพลิดเพลินจากการเรียน

แต่อย่างไรก็ตามจากการสำรวจพบว่านักเรียนขาดความรู้ พื้นฐานด้านภาษาอังกฤษตามแผนการสอน ขาดความสนใจในการ เรียนรู้วิชาภาษาอังกฤษ และ ขาดความมั่นใจในการ แสดงออกด้านภาษาอังกฤษ ทั้งนี้ผลสรุปดังกล่าวสรุปมาจากที่ ผู้วิจัยทำการทดสอบนักเรียนด้วยการสัมภาษณ์และสอบถาม เป็นรายบุคคล ดังนั้นผู้วิจัยจึงออกแบบคำศัพท์ที่จะนำมาใส่ไว้ ในเกม จำนวน 150 คำ โดยแบ่งเป็นกลุ่มของคำศัพท์ อัน

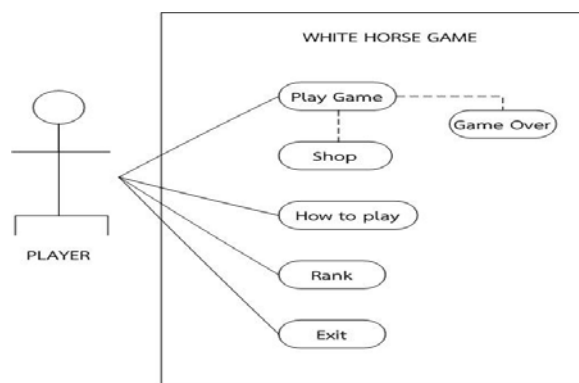
ประกอบด้วย ตัวเลข สี วัน เดือน อารมณ์ อวัยวะในร่างกาย ญาติ คำกริยา ศัพท์ในห้องเรียน สัตว์ และผลไม้

4.2 รูปแบบของเกม

ผู้พัฒนาได้นำรูปแบบเกมการผจญภัยเข้ามาใช้เพื่อช่วยใน การดึงดูดความสนใจให้แก่ผู้เล่นโดยแทรกเนื้อหาการเรียนรู้อังกฤษในด้านต่างๆเข้าไปในเกมโดยผู้เล่นจะบังคับม้า เพื่อจะผจญภัยไปในด้านต่างๆจำนวน 5 ด้าน เพื่อเก็บแครอท และต้องหลบอุปสรรคเช่นหนาม หิน เป็นต้น จากนั้นจะพบด่านทดสอบพื้นฐานความรู้ภาษาอังกฤษในด้านการอ่าน การฟัง การพูด การเขียน และการสะกดคำภาษาอังกฤษตามแต่ละด่าน (ดังแสดงในภาพที่ 3 - 8) ซึ่งเป็นกระบวนการเรียนรู้หลักที่ ต้องการให้ผู้เล่นได้รับโดยตรงจากการเล่นเกม White Horse ซึ่ง เป็นการสร้างแรงจูงใจให้ผู้เล่นเกิดความพยายามที่จะผ่าน อุปสรรคทั้งจากตัวเกมและตัวเนื้อหาที่แทรกเข้าไปในเกมจน เกิดเป็นกระบวนการเรียนรู้โดยไม่รู้ตัว

4.3 การออกแบบและพัฒนาเกม White Horse

การออกแบบกิจกรรมการเล่นเกมผู้วิจัยอาศัยแนวคิด ของ Activity Diagram ดังแสดงตัวอย่างในรูปที่ 1 และด้านการ พัฒนา ผู้วิจัย ใช้ JDK (Java Development Kit), Android Development Tools (ADT) plugin เป็นเทคโนโลยีในการ พัฒนาเกมและโปรแกรม PaintStudio ในการวาดภาพประกอบ

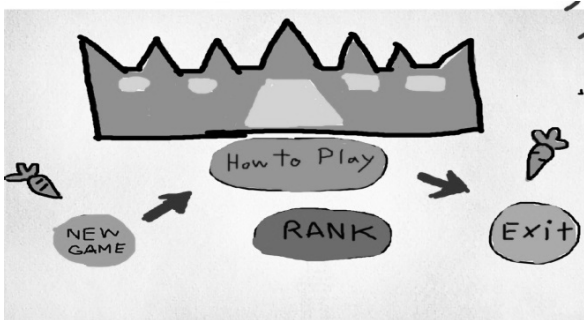


ภาพที่ 1: Activity Diagram การออกแบบเกม White Horse

ผู้วิจัยออกแบบตามทฤษฎีการสร้างเกมโดยแบ่งเกมออกเป็น 2 ส่วน คือส่วนที่เป็นเกมเพื่อความสนุกสนานดึงดูดความสนใจผู้เล่น และส่วนที่เป็นกระบวนการเรียนรู้ซึ่งแทรกเนื้อหาทักษะ ภาษาอังกฤษเข้าไป ดังรายละเอียดต่อไปนี้

เมื่อผู้เล่นเปิด โปรแกรมขึ้นมาจะปรากฏหน้าจอเมนูหลักซึ่งจะมีพื้นหลังเป็นสีเขียว โดยปริมาตรเป็นฉากหลังของเกม โดยเมนูจะถูกจัดเรียงทางด้านล่างของรูปปริมาตรซึ่งมีปุ่มอยู่ 4 แบบ คือ New Game , How to Play , Rank และ Exit ดังแสดงในภาพที่ 2 แต่ละปุ่มจะมีลักษณะเป็นวงรีล้อมตัวอักษรเพื่อต่อการใช้งาน และมีลูกศรชี้กำกับเพื่อช่วยเสริมความเข้าใจในการเล่น

ส่วนภาพที่ 3 เป็นหน้าจอหลักของเกม เป็นการเข้าสู่สาระเรียนรู้ที่ทักษะภาษาอังกฤษ โดยแต่ละด้านต้องผ่านการเล่นเกมมาก่อน ผู้เล่นจะต้องบังคับม้าด้วยการใช้ปุ่มบังคับด้านซ้ายมือของจอ เพื่อเก็บแครอทสะสมพลังและคะแนนสำหรับอัปเกรดในด้านถัดไป ระหว่างเก็บแครอทผู้เล่นจะพบกับอุปสรรค เช่น งู ดอกไม้พิษ หิน ฉากที่ต้องเล่นมีทั้งหมด 5 ฉากหรือ 5 ด้าน



ภาพที่ 2: หน้าจอหลักของเกม White Horse



ภาพที่ 3: ฉากการเล่นในเกมในส่วนผจญภัย

ภาพที่ 4 เป็นส่วนของเกมเพื่อเสริมการเรียนรู้ เมื่อผู้เล่นเก็บแครอทจนครบตามที่กำหนด หน้าจอจะแสดงฉากฝึกทักษะภาษาอังกฤษ การฟัง ผู้เล่นต้องคลิกที่ภาพเพื่อฟังเสียงอ่านของคำศัพท์ ผู้เล่นต้องเลือกภาพให้ครบจำนวน 5 ภาพ เพื่อฟังคำศัพท์ 5 คำจากนั้นจึงจะผ่านไปด้านต่อไปได้



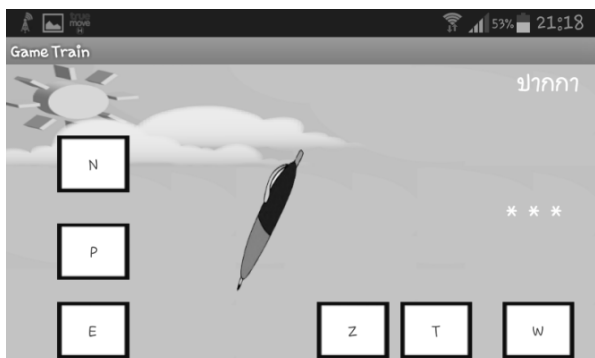
ภาพที่ 4: ฉากเล่นเกมเพื่อฝึกทักษะการฟัง



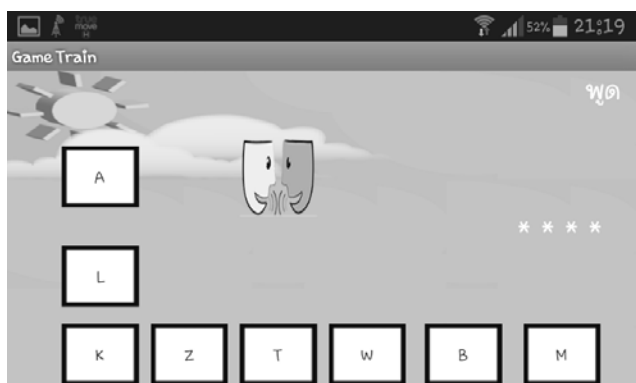
ภาพที่ 5: ฉากเล่นเกมเพื่อฝึกทักษะการพูดและการอ่าน

จากภาพที่ 5 เป็นการแสดงการเล่นเพื่อฝึกทักษะการพูดและการอ่านซึ่งจะปรากฏในเกมด้านที่ 2 และด้านที่ 3 ผู้วิจัยได้ใช้ตัวจับการออกเสียงของ Google โดยผู้เล่นต้องคลิกที่คำศัพท์ จากนั้นจะได้ยินเสียงอ่านของคำศัพท์ดังขึ้น จากนั้นผู้เล่นจะต้องทำการคลิกที่รูปไมโครโฟนเพื่อออกเสียงตามจนครบทั้ง 5 คำ ถึงจะผ่านไปเล่นในฉากต่อไปได้ ข้อแตกต่างระหว่างด้าน 2 และด้าน 3 คือระบบจะจับเสียงในด้าน 3 ละเอียดกว่าในด้านที่ 2

จากภาพที่ 6 เป็นการแสดงหน้าจอเกมด้านที่ 4 ผู้เล่นต้องคลิกที่รูปภาพที่ปรากฏ เพื่อฟังเสียงอ่านของคำศัพท์ ดังเช่นตัวอย่างที่ปรากฏในภาพคือรูปปากกา (PEN) หลังจากนั้นเลือกคลิกกล่องพยัญชนะที่ตรงกับคำศัพท์ตามรูปภาพ เช่น คลิกที่กล่องตัว P ตัวอักษร P จะไปปรากฏแทนที่รูปดอกจันที่อยู่ขวามือตามตำแหน่งของ P ภายในคำว่า PEN แต่หากคลิกเลือกตัวอักษรผิด เช่น คลิกตัวอักษร Z ตัวอักษร Z จะไม่ปรากฏที่กลุ่มดอกจันทางขวามือ กรณีนี้เกมจะจบลงเมื่อผู้เล่นคลิกต่อคำศัพท์จนครบทุกคำ นั้นหมายความว่าด้านนี้จะเป็นการฝึกทักษะการฟังและการสะกดคำ



ภาพที่ 6: ฉากเล่นเกมเพื่อฝึกทักษะการสะกดคำ



ภาพที่ 7: ฉากเล่นเกมเพื่อฝึกทักษะการเขียนและการฟัง

จากภาพที่ 7 เป็นการแสดงหน้าจอเกมด่านที่ 5 ผู้เล่นต้องคลิกที่รูปภาพที่ปรากฏเพื่อฟังเสียงอ่านคำศัพท์ ดังตัวอย่างคือ รูปพูดคุย (TALK) จากนั้นคลิกกล่องพยัญชนะที่ตรงกับคำศัพท์ของภาพ เช่น คลิกกล่องตัว T แล้วตัวอักษร T จะไปปรากฏแทนที่รูปดอกจันที่อยู่ขวามือตามตำแหน่งของ T แต่หากคลิกเลือกตัวอักษรผิดจะไม่ปรากฏที่กลุ่มดอกจันทางขวามือ และเกมจะจบลงเมื่อผู้เล่นคลิกต่อคำศัพท์จนเสร็จ ด่านนี้จะเป็นการฝึกทักษะการฟังและการเขียน

4.4 ผลการทดลอง

ผู้วิจัยนำแบบสอบถามจากงานวิจัยของอัปสร อีซอ [13] มาใช้เป็นต้นแบบ เพื่อการศึกษาค่าความพึงพอใจของนักเรียนต่อเกม White Horse ผลการเก็บข้อมูลจากแบบสอบถาม พบว่านักเรียนชั้น ป.1 มีค่าความพึงพอใจด้วยค่าเฉลี่ยเท่ากับ 5.00 (S.D.=0.00) ซึ่งอยู่ในระดับ พึงพอใจมากที่สุด ส่วน ป.2 มีค่าความพึงพอใจด้วยค่าเฉลี่ยเท่ากับ 4.90 (S.D.=0.23) อยู่ในระดับ พึงพอใจมากที่สุด และ ป.3 มีค่าความพึงพอใจด้วยค่าเฉลี่ยเท่ากับ 4.42 (S.D.=0.65) อยู่ในระดับ พึงพอใจมากที่สุด

นอกจากนี้ผู้วิจัยได้ทำการทดสอบผลสัมฤทธิ์ของการเล่นเกม เพื่อศึกษาผลการสอดแทรกความรู้ด้านศัพท์ภาษาอังกฤษผ่านเกม ผลการทดสอบด้วยแบบทดสอบทักษะการเขียน การสะกด การพูด จำนวน 3 ชุด จากการสุ่มเลือกแบบทดสอบทั้งก่อนเล่นและหลังเล่น ผลการทดสอบผลสัมฤทธิ์การเล่นเกมนักเรียนชั้น ป.1 – 3 ดังแสดงในภาพที่ 8 – 10 ตามลำดับ

คนที่	เขียน สะกด พูด			ฟัง อ่าน		
	ก่อนเล่น	หลังเล่น	ผลลัพท์	ก่อนเล่น	หลังเล่น	ผลลัพท์
1	0	5	5	6	10	4
2	0	5	5	4	11	7
3	1	5	4	3	11	8
4	1	6	5	10	13	3
5	0	5	5	12	10	-2
x bar	0.40	5.20	4.80	7.00	11.00	4.00
s.d.	0.55	0.45	0.45	3.87	1.22	3.94

ภาพที่ 8: ผลการทดสอบก่อนเล่นและหลังเล่นของชั้น ป.1 จำนวน 5 คน พบว่านักเรียนมีพัฒนาการด้านทักษะการฟังและทักษะการอ่านสูงขึ้นเฉลี่ย 4.00 (S.D. = 3.94)

คนที่	เขียน สะกด พูด			ฟัง อ่าน		
	ก่อนเล่น	หลังเล่น	ผลลัพท์	ก่อนเล่น	หลังเล่น	ผลลัพท์
1	4	16	12	6	18	12
2	3	12	9	12	20	8
3	3	11	8	18	20	2
4	4	13	9	18	22	4
5	1	11	10	16	20	4
6	1	8	7	20	20	0
7	0	7	7	18	24	6
x bar	2.29	11.14	8.86	15.43	20.57	5.14
s.d.	1.60	3.02	1.77	4.86	1.90	3.98

ภาพที่ 9: ผลการทดสอบก่อนเล่นและหลังเล่นของชั้น ป.2 จำนวน 7 คน พบว่านักเรียนมีพัฒนาการด้านทักษะการฟังและทักษะการอ่านสูงขึ้นเฉลี่ย 5.14 (S.D. = 3.98)

คนที่	เขียน สะกด พูด			ฟัง อ่าน		
	ก่อนเล่น	หลังเล่น	ผลลัพท์	ก่อนเล่น	หลังเล่น	ผลลัพท์
1	0	21	21	16	22	6
2	6	7	1	6	18	12
3	3	4	1	20	24	4
4	7	8	1	3	16	13
5	0	9	9	20	24	4
6	3	9	6	22	22	0
7	3	21	18	20	26	6
8	3	18	15	24	28	4
9	4	9	5	20	26	6
10	5	6	1	24	26	2
11	4	6	2	24	26	2
12	7	19	12	30	30	0
x bar	3.75	11.42	7.67	19.08	24.00	4.92
s.d.	2.30	6.37	7.25	7.65	4.00	4.12

ภาพที่ 10: ผลการทดสอบก่อนเล่นและหลังเล่นของชั้น ป.3 จำนวน 12 คน พบว่า นักเรียนมีพัฒนาการด้านทักษะการฟังและทักษะการอ่านสูงขึ้นเฉลี่ย 4.92 (S.D. = 4.12)

5. สรุป

รายงานวิจัยฉบับนี้นำเสนอเกม White Horse บนระบบปฏิบัติการแอนดรอยด์ ที่มีการสอดแทรกคำศัพท์ภาษาอังกฤษระดับชั้นประถมศึกษาปีที่ 1-3 จำนวน 150 คำ ในแต่ละด่านของเกม จำนวน 5 ด่าน ทั้งนี้เพื่อเป็นการฝึกทักษะการเขียน การสะกดคำ การพูด การฟัง และการอ่าน โดยที่ผู้เล่นจะเกิดการเรียนรู้คำศัพท์ภาษาอังกฤษแบบไม่รู้ตัว

การเก็บผลทดลองกับนักเรียนชั้นประถมศึกษาปีที่ 1-3 โรงเรียนบ้านศาลา ตำบลคงใหญ่ อำเภอยะหา จังหวัดนครราชสีมา จำนวน 24 คน พบว่า ผู้เล่นมีความพึงพอใจต่อเกมในระดับพึงพอใจมาก ด้วยค่าเฉลี่ยเท่ากับ 5.00 (S.D.=0.00), 4.90 (S.D.=0.23), และ 4.42 (S.D.=0.65) ตามลำดับ นอกจากนี้ การทดลองยังพบว่า นักเรียนชั้นประถมศึกษาปีที่ 1-3 มีพัฒนาการด้านการเขียน การสะกด การพูด ด้วยค่าเฉลี่ยเท่ากับ 4.80 (S.D.=0.45), 8.86 (S.D.=1.77), และ 7.67 (S.D.=7.25) ตามลำดับ อีกทั้งพัฒนาการด้านการฟังและการอ่าน มีค่าเฉลี่ยเท่ากับ 4.00 (S.D.=3.94), 5.14 (S.D.=3.98), และ 4.92 (S.D.=4.12) ตามลำดับ

จากผลการทดลองด้านความพึงพอใจต่อเกมและผลการพัฒนาทักษะภาษาอังกฤษของนักเรียนชั้นประถมศึกษาปีที่ 1-3 สามารถสรุปผลได้ว่า เกม White Horse สามารถนำไปใช้เล่นเกมเสริมการเรียนรู้ของนักเรียนชั้นประถมศึกษาปีที่ 1-3 ได้ พร้อมทั้งยังส่งเสริมการเรียนรู้นอกชั้นเรียนได้อย่างมีประสิทธิภาพ

อย่างไรก็ตาม เนื่องจากเกมนี้เป็นเพียงเกมต้นแบบในการพัฒนาเกมเสริมทักษะภาษาอังกฤษให้กับนักเรียนที่อยู่ห่างไกลชุมชนเมือง การจัดการเรียนการสอนวิชาภาษาอังกฤษในชั้นเรียนยังคงมีความสำคัญ เกมเป็นเพียงตัวเสริมในการสร้างความมั่นใจและการเรียนรู้ของนักเรียนนอกชั้นเรียนเท่านั้น

เอกสารอ้างอิง

- [1] สำนักงานเลขาธิการสภาผู้แทนราษฎร, *การก้าวสู่ประชาคมอาเซียน*, สำนักพิมพ์ สำนักงานเลขาธิการสภาผู้แทนราษฎร: กรุงเทพมหานคร, 2555.
- [2] ไพฑูริย์ ศรีฟ้า, “เปิดโลก Tablet สู่ทิศทางการวิจัยด้านเทคโนโลยีและสื่อสารการศึกษา: จากแนวคิดสู่กระบวนการปฏิบัติ” *ภาควิชาเทคโนโลยีการศึกษา มหาวิทยาลัยเกษตรศาสตร์ มหาวิทยาลัยทักษิณ วิทยาเขตสงขลา*, วันที่ 4 กันยายน, สงขลา, 2554.
- [3] สุรศักดิ์ ปาเฮ. “แท็บเล็ตเพื่อการศึกษา: โอกาสและความท้าทาย” <http://www.addkute3.com>, ค้นเมื่อ 2 กันยายน, 2555.
- [4] ชูวิทย์ มิตรชอบ, “ยุทธศาสตร์การเข้าสู่ประชาคมอาเซียน ปี พ.ศ. 2558” *สำนักงานคณะกรรมการพัฒนาการเศรษฐกิจและสังคมแห่งชาติ หอประชุมสุมนัยประดิษฐ์ สำนักงาน ก.พ.*, 17 ธันวาคม 2555, นนทบุรี, 2555.
- [5] จริญญา สอนสุด, *ผลการสอนโดยใช้เกมที่มีต่อความสามารถในการเรียนรู้คำศัพท์ ความคงทนในการเรียนรู้ และเจตคติต่อการเรียนวิชา*

ภาษาอังกฤษของนักเรียนชั้นประถมศึกษาปีที่ 2, วิทยานิพนธ์มหาบัณฑิต สาขาหลักสูตรและการสอน แขนงภาษาอังกฤษ มหาวิทยาลัยราชภัฏนครสวรรค์: นครสวรรค์, 2550.

- [6] I. Sriprasidh, “New Research-based Literacy Framework in ELT in Thailand by integrating the Phonics approach (bottom-up) with the Whole Language approach (top-down) in teaching Thai to read & write English as well as to overcome “dyslexia”, *UNESCO-APEID International Conference the 4th World Teachers' Day and 12th in Thailand*, 2009.
- [7] B. Punyapet and C. Laohawiriyanon, “The Effects of Remedial English Class Using Systemic Phonics to Improve Students' Pronunciation, Spelling and Reading Comprehension Proceedings-Teaching Techniques-005”, *Faculty of Liberal Arts, Prince of Songkla University, 4th International Conference on Humanities and Social Sciences*, April 21st, 2012, Songkla, 2012.
- [8] E. Klopfer, S. Osterweil, and K. Salen, “Moving learning games forward,” *The Education Arcade Massachusetts Institute of Technology*, USA, 2009.
- [9] นายวิสิทธิ์ กาญจนธม, “เกมบนหน้าจอสัมผัสบนระบบปฏิบัติการแอนดรอยด์ (Android Game via Touchscreen)”, *โครงการระดับปริญญาตรี สาขาวิชาวิทยาการคอมพิวเตอร์ ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย* 2555.
- [10] Z. Muda and I. S. Basiron, "Multimedia Adventure Game As Edutainment Application ", *Proceedings of the 2005 International Conference on Computational Intelligence for Modelling, Control and Automation, and International Conference on Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-LAWTIC'05)*. Vol. 2, pp. 497 – 500, November 2005.
- [11] K. Ali and Hossain, M. Anwar and E. Saddik, Abdulmotaleb and G. Wail, "A Multimedia-driven Ambient Edutainment System for the Young Children", *Proceeding of ACM international workshop on Story representation, mechanism and context (SRMC'08)*, Vol. 2, pp 57-64, ACM New York, 2008.
- [12] G. Veenhof, "ZooQuest: A mobile game-based learning application for fifth graders", *Thesis Master Informatiekunde Programma Human Centered Multimedia*, University of Amsterdam, 2011.
- [13] อัสพร อีซอ, *การพัฒนารูปแบบการสอนแบบเกมในการจัดการเรียนการสอนสายวิชาการตลาด(ศึกษาเฉพาะกรณีมหาวิทยาลัยราชภัฏกลุ่มภูมิศาสตร์ภาคใต้)*, รายงานวิจัย คณะวิทยาการจัดการ มหาวิทยาลัยราชภัฏยะลา: ยะลา, 2549-2550.

การรู้จำท่ามือภาษาไทยโดยใช้ค่าความถี่ของทิศทางตามค่าเกรเดียนท์ร่วมกับ การวิเคราะห์องค์ประกอบหลัก และซัพพอร์ตเวกเตอร์แมชชีน Thai Hand Shape Recognition Using HOG - PCA and SVM

อนุสรณ์ อุ่นท้าว (Anusorn Untao)¹ และสมปอง เวฬุวนาธร (Sompong Valuvanathon)²

ภาควิชา คณิตศาสตร์ สถิติ และคอมพิวเตอร์ คณะวิทยาศาสตร์

มหาวิทยาลัยอุบลราชธานี

¹amatadepp@gamil.com, ²scsompva@ubu.ac.th

บทคัดย่อ

บทความนี้มีวัตถุประสงค์เพื่อนำเสนอขั้นตอนวิธีการเพิ่มประสิทธิภาพการรู้จำท่ามือภาษาไทย โดยใช้ค่าความถี่ของทิศทางตามค่าเกรเดียนท์ (Histograms of Oriented Gradients Algorithm: HOG) ในการสกัดคุณลักษณะของรูปภาพท่ามือในภาษาไทย จากนั้นใช้การวิเคราะห์องค์ประกอบหลัก (Principal component analysis: PCA) ทำการแปลงโครงสร้างเมตริกซ์คำนวณหาไอเกนเวกเตอร์ที่สอดคล้องกันกับค่าไอเกน เลือกเอาเฉพาะส่วนที่เป็นนัยสำคัญเพื่อลดขนาดมิติของข้อมูล เพื่อนำไปประมวลผลการรู้จำด้วยวิธีการ ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine: SVM) แบบเรเดียลเบสฟังก์ชันเคอร์เนล (Radial Basis Function : RBF) โดยทดลองกับชุดข้อมูลรูปถ่ายท่ามือภาษาไทยที่จัดทำขึ้นจำนวน 710 ภาพ ซึ่งประกอบด้วยท่ามือจากผู้แสดงจำนวน 5 คน คนละ 142 ท่า ผลการทดลองแสดงให้เห็นว่าวิธีการที่นำเสนอสามารถเพิ่มประสิทธิภาพการรู้จำได้โดยมีค่าความถูกต้องในการรู้จำ สูงสุดร้อยละ 96

คำสำคัญ: ภาษามือไทย การรู้จำท่ามือ ค่าความถี่ของทิศทางตามค่าเกรเดียนท์ การวิเคราะห์องค์ประกอบหลัก ซัพพอร์ตเวกเตอร์แมชชีน เรเดียลเบสฟังก์ชันเคอร์เนล

Abstract

This paper presents a technique to enhance the accuracy of Thai hand shape recognition. Histograms of oriented gradients (HOG) was used to extract features while principal component analysis (PCA) was used to reduce feature dimensionality. In the part of recognition,

the support vector machine (SVM) with radial basis function (RBF) kernel was used to classify the features. The experiment dataset consists of 710 images, 142 hand shapes from five individuals. The experimental results show that the proposed method can enhance the performance of Thai hand shape recognition up to 96%.

Keyword: Thai sign, Hand Shape Recognition, HOG, PCA, SVM, RBF.

1. บทนำ

ภาษามือไทย (Thai Sign Language) เป็นภาษาสื่อสารสำหรับผู้พิการทางหูในประเทศไทย ซึ่งได้ถูกรับรองให้เป็นภาษามือประจำชาติไทย [1] ปัจจุบันมีการศึกษาวิจัยเกี่ยวกับการประยุกต์และพัฒนาเทคโนโลยีด้านรู้จำท่ามือทั้งในประเทศและนานาชาติ โดยอิงหลักพื้นฐานของการมีปฏิสัมพันธ์ระหว่างมนุษย์กับคอมพิวเตอร์ (Human Computer Interaction: HCI) ถ้าหากต้องการให้ผู้พิการทางหูสามารถสื่อสารกับบุคคลทั่วไปผ่านคอมพิวเตอร์ได้อย่างมีประสิทธิภาพ การรู้จำท่ามือจำเป็นต้องมีความถูกต้องและแม่นยำ [2] ซึ่งองค์ประกอบของภาษามือประกอบด้วย ท่ามือ ตำแหน่ง การพลิกหันฝ่ามือ การเคลื่อนไหว และสีหน้า โดยท่ามือถือว่าเป็นหัวใจสำคัญที่ใช้สื่อสารสำหรับผู้พิการทางหู ในภาษามือไทยมีท่ามือจำนวนทั้งหมด 142 แบบที่แตกต่างกัน [3],[4] จากการศึกษางานวิจัยที่ผ่านมาพบว่าชุดข้อมูล ที่ใช้ทดสอบยังไม่ครอบคลุมท่ามือทั้ง 142 แบบ [5], [6], [7], [8], [9] และยังไม่สามารถแก้ปัญหาในส่วนของการจำแนกท่ามือที่มี

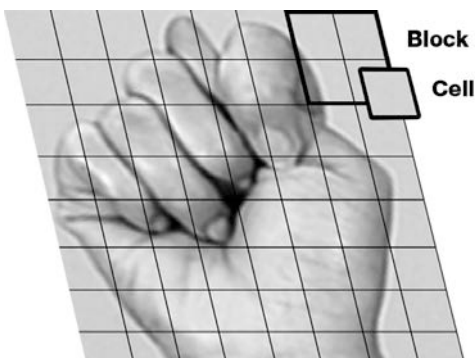
ลักษณะคล้ายคลึงกันอย่างมากจนทำให้เกิดความกำกวมของ
ภาษามือ และการประมวลผลที่ผิดพลาดในภาษามือไทย [10]

จากปัญหาดังกล่าวผู้วิจัยจึงมีแนวคิดในการพัฒนาขั้นตอน
วิธีการสำหรับการรู้จำท่ามือภาษามือไทยทั้ง 142 แบบ โดยใช้
ความถี่ของทิศทางตามค่าเกรเดียนท์ร่วมกับการวิเคราะห์
องค์ประกอบหลักในขั้นตอนการสกัดคุณลักษณะ และใช้การ
จำแนกข้อมูลด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีน แบบเบเรเดิล
เบสิสฟังก์ชันเคอร์เนล โดยมีวัตถุประสงค์เพื่อนำเสนอขั้นตอน
วิธีการเพิ่มประสิทธิภาพการรู้จำที่แม่นยำ และการลดระยะเวลา
ในการประมวลผล

2. วรรณกรรมที่เกี่ยวข้อง

2.1 ค่าความถี่ของทิศทางตามค่าเกรเดียนท์ (Histograms of Oriented Gradients: HOG)

การดึงลักษณะเด่นของท่ามือโดยใช้คำอธิบายภาพแบบ
ค่าความถี่ของทิศทางตามค่าเกรเดียนท์ เป็นวิธีการดึงลักษณะ
รูปร่างภายในภาพโดยใช้การกระจายตัวของความเข้มเกร
เดียนท์ หรือทิศทางของเส้นขอบ โดยใช้วิธีการแบ่งภาพ
ออกเป็นเซลล์ (Cell) เล็กๆ ในแต่ละเซลล์นั้นจะประกอบด้วย
ทิศทางค่าเกรเดียนท์ ซึ่งถูกเก็บไว้ในรูปแบบของฮิสโตแกรม
ที่อธิบายคุณลักษณะวัตถุที่อยู่ในเซลล์ เพื่อเพิ่มประสิทธิภาพของ
ความถูกต้องสามารถนำฮิสโตแกรมมาทำนอร์มอลไลซ์ ด้วย
การคำนวณตัวชี้วัดค่าความเข้มจากโอเวอร์แลปของเซลล์ภายใน
บล็อก (Block) เพื่อลดผลกระทบจากการเปลี่ยนแปลงของแสง
และเงาให้น้อยลง



ภาพที่ 1: ตัวอย่างการแบ่งภาพเป็น เซลล์และบล็อก

การประมวลผลค่าเกรเดียนท์สามารถหาได้จากการใช้
1D - discrete derivative masks หากำอนุพันธ์อันดับหนึ่งตาม
แนวตั้งและแนวนอนด้วยวิธีการ $[-1, 0, 1]$ and $[-1, 0, 1]^T$ ใน
การกำหนดค่า mask ของแกน x และแกน y เพื่อนำมาใช้ในการ
การคำนวณสร้างเซลล์ฮิสโตแกรมที่มีความสัมพันธ์กัน โดยแต่
ละพิกเซลที่อยู่ในเซลล์จะมีทิศทางและน้ำหนักจากการ
คำนวณหาค่าเกรเดียนท์ใกล้เคียงกันทำให้เซลล์ที่อยู่ในกลุ่ม
เดียวกันจะมีค่าฮิสโตแกรมที่คล้ายกัน เมื่อผ่านการคำนวณหา
ค่าเกรเดียนท์และการหาทิศทางของภาพแล้วจะได้คำอธิบาย
ภาพในรูปแบบฮิสโตแกรม

2.2 การวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis: PCA)

การวิเคราะห์องค์ประกอบหลัก เป็นวิธีการที่ใช้ในทฤษฎี
การแทนข้อมูลลงในเวกเตอร์ ด้วยการแปลงโครงสร้าง
เมตริกซ์เพื่อทำการลดขนาดของข้อมูล โดย ด้วยการคำนวณหา
ไอเกนเวกเตอร์ที่สอดคล้องกันเพื่อเลือกเอาเฉพาะข้อมูลที่มี
นัยสำคัญ โดยนำภาพขนาด $N \times M$ แปลงเป็นเวกเตอร์ภาพ
เรียนรู้เท่ากับ L ภาพ และชุดของเวกเตอร์ภาพเรียนรู้คือ
 $\{X_i ; i=0, \dots, L-1\}$ โดยสมมุติว่าเวกเตอร์มีขนาดเท่ากับขนาด
ของภาพ หลังจากนั้นทำการคำนวณหาค่าเฉลี่ยของภาพได้ดัง
สมการ

$$x_m = \frac{1}{L} \sum_{i=0}^{L-1} X_i \quad (1)$$

และคำนวณหาค่า Covariance Matrix ดังสมการ

$$C = \frac{1}{L} \sum_{i=0}^{L-1} (X_i - x_m)^T (X_i - x_m) \quad (2)$$

หลังจากนั้นทำการวิเคราะห์แยกส่วนค่าไอเกน (Eigenvalue
decomposition) กับเมตริกซ์ C ดังสมการ

$$C = U \Lambda U^T = \sum_{i=0}^{L-1} \lambda_i u_i u_i^T \quad (3)$$

โดยที่ λ_i คือ Eigenvalue ของเมตริกซ์ C โดย Eigenvalue นั้น
จะเรียงจากมากไปหาน้อย และ u_i คือ Eigenvector ซึ่งชุด

Eigenvector u_i นี้เรียกว่า Eigenfaces คุณสมบัติที่สำคัญของการวิเคราะห์แยกส่วนที่ได้จะเป็นค่าสัมประสิทธิ์ที่เหมาะสม (Optimal Coefficients) a_i สำหรับ X คือค่าที่ต้องการนำมาวิเคราะห์ให้ใด ๆ ซึ่งค่า Coefficients สามารถคำนวณหาได้จากสมการ

$$a_i = (X - X_m) \cdot u_i \quad (4)$$

2.3 การจำแนกข้อมูลแบบซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine: SVM)

ซัพพอร์ตเวกเตอร์แมชชีน ได้ถูกนำมาใช้ในการจำแนกด้วยสมการจำแนกค่าคุณลักษณะของสองกลุ่มที่วางตัวอยู่ในพื้นที่คุณลักษณะ (Feature Space) ออกจากกันโดยการสร้างเส้นแบ่งที่เป็นเส้นตรงเพื่อให้ทราบว่าเส้นตรงที่แบ่งข้อมูลออกจากกันเส้นตรงใดเป็นเส้นตรงที่ดีที่สุด สามารถเขียนเป็นสมการเพื่อใช้ในการแก้ปัญหา โดยนำข้อมูลมาวางไว้ในพื้นที่คุณลักษณะกลุ่มข้อมูลในรูปแบบเวกเตอร์จะได้

$$X = ((x_1 - y_1), \dots, (x_i - y_i)) \quad (5)$$

กำหนดสมการตัวบ่งชี้ข้อมูลเส้นแบ่งไฮเปอร์เพลน โดยเส้นขอบ (Margin) ของเส้นแบ่งเป็นเส้นที่สัมผัสกับค่าข้อมูลในพื้นที่คุณลักษณะที่ใกล้สุดเส้นขอบของเส้นทั้งสองถูกแทนด้วย $w^T x + b \geq y \geq 1$ ถ้า $y = 1$ และ $w^T x + b \leq y \leq -1$ ถ้า $y = -1$ หากเส้นขอบใดๆ ที่มีค่ากว้างมากที่สุดจะแสดงให้เห็นข้อมูลที่แยกกันมากที่สุด ดังนั้นเส้นขอบที่มีความกว้างมากที่สุดจะเป็นเส้นแบ่งข้อมูลที่ดีที่สุด และเรียกตำแหน่งสัมผัสข้อมูลที่ใกล้ที่สุดจากเส้นขอบว่า เวกเตอร์สนับสนุน (Support Vector) โดย เรียกเส้นแบ่งข้อมูลทั้งสองว่าเส้นขอบ ซึ่งสามารถคำนวณหาความกว้างของเส้นขอบในรูปแบบปกติ (Normalization) โดยคำนวณจากสมการ

$$\begin{aligned} w^T x^+ + b &= 1 \\ w^T x^- + b &= -1 \\ w^T (x^+ + x^-) &= 2 \\ M &= \left(\frac{w}{\|w\|} \right)^T (x^+ + x^-) \\ M &= \frac{2}{\|w\|} \end{aligned} \quad (6)$$

แต่ในทางปฏิบัตินั้นข้อมูลอาจจับกลุ่มกันในตำแหน่งต่าง ๆ ทำให้ไม่สามารถใช้สมการซัพพอร์ตเวกเตอร์แมชชีนเชิงเส้นได้ วิธีการเคอร์เนล (Kernel) ถูกนำมาใช้แก้ปัญหา โดยการแปลงมิติของข้อมูลให้สูงขึ้นเพื่อช่วยในการจัดเรียงตัวของข้อมูลที่ใหม่ ที่เรียกว่า พื้นที่มิติสูง (Higher Dimensional Space) ซึ่งเป็นวิธีการที่ซับซ้อนจะต้องอาศัยความรู้และความชำนาญในการปรับค่าพารามิเตอร์เพื่อให้ได้ประสิทธิภาพในการจำแนกที่ดีที่สุด ซึ่งปัจจุบันเคอร์เนลที่นิยมใช้มีอยู่ 3 แบบด้วยกันคือ

1) โพลีโนเมียล (Polynomial)

$$k(x_i, x_j) = (y \cdot x_i^T x_j + r)^d, y > 0 \quad (7)$$

2) เรเดียลเบสิสฟังก์ชัน (Radial Basis Function)

$$k(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0 \quad (8)$$

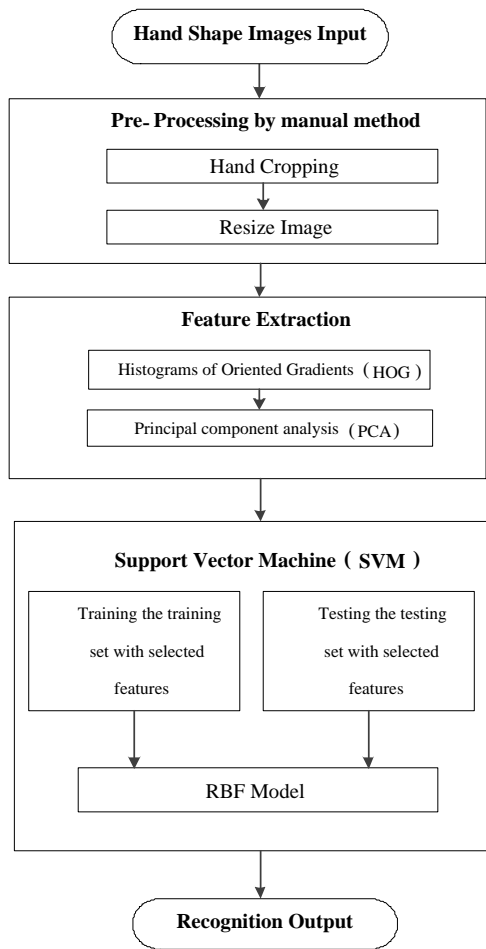
3) ซิกมอยด์ (Sigmoid)

$$k(x_i, x_j) = \tan(y x_i^T x_j + r) \quad (9)$$

ซึ่งในงานวิจัยนี้ เลือกใช้ เรเดียลเบสิสฟังก์ชันเคอร์เนล ซึ่งผู้ทดลองได้ทดสอบการจำแนกทดสอบกับชุดข้อมูลเบื้องต้นพบว่า เรเดียลเบสิสฟังก์ชันเคอร์เนลให้ผลการรู้จำสูงสุด ซึ่งสอดคล้องกับข้อบ่งชี้ในการเพิ่มประสิทธิภาพการจำแนกข้อมูลด้วยวิธี ซัพพอร์ตเวกเตอร์แมชชีนด้วยเรเดียลเบสิสฟังก์ชันเคอร์เนล [13], [14]

3. วิธีการทดลอง

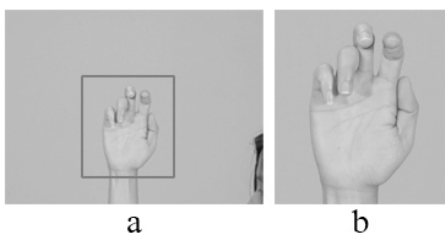
การทดลองสามารถแบ่งขั้นตอนเป็น 3 ขั้นตอนหลักดังนี้ การจัดเตรียมข้อมูล (Preprocessing) การสกัดคุณลักษณะ (Feature Extraction) และการจำแนกข้อมูล (Classification) โดยทดลองกับชุดข้อมูล รูปถ่ายท่ามือภาษาไทยที่จัดทำขึ้นในสถานะแสงคงที่มีพื้นหลังที่ไม่ซับซ้อนจำนวน 710 ภาพ ซึ่งประกอบด้วยท่ามือในภาษาไทยที่ถูกรวบรวมไว้ทั้งหมดจำนวน 142 แบบ ใช้ผู้แสดงท่ามือจำนวน 5 คน และการจัดทำชุดข้อมูลผ่านการตรวจสอบความถูกต้องจากผู้เชี่ยวชาญทางด้านภาษาไทยจำนวน 3 คน



ภาพที่ 2: ฟังแสดงโครงสร้างการออกแบบการทดลอง

3.1 ขั้นตอนการจัดเตรียมข้อมูล

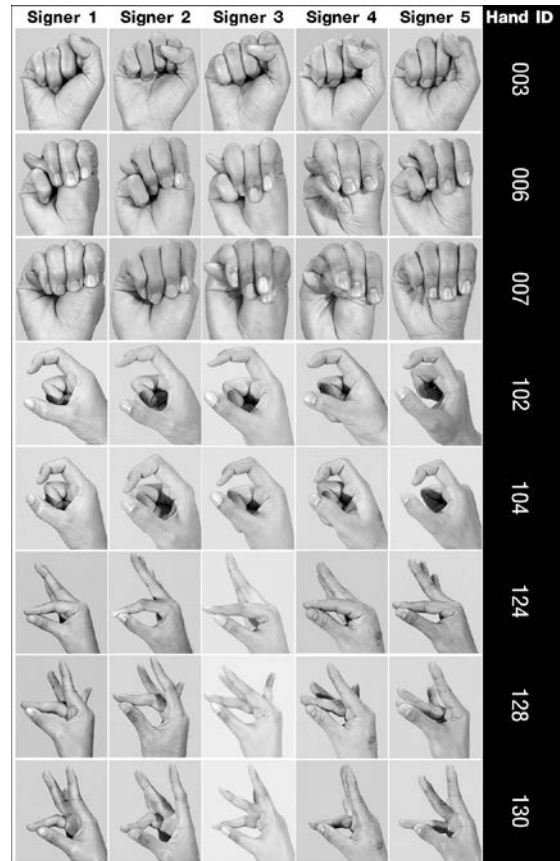
กระบวนการจัดเตรียมข้อมูลในการทดลองนี้ได้ใช้วิธีการทำมือ (Manual Method) โดยทำการเลือกตัดเอาเฉพาะบริเวณที่เป็นมือจากภาพถ่ายขนาด $4,608 \times 3,456$ (a) แล้วทำการย่อตามสัดส่วนให้ได้ภาพขนาด 180×200 (b) แสดงในภาพที่ 3



ภาพที่ 3: ภาพตัวอย่างที่ผ่านกระบวนการการจัดเตรียมข้อมูล

3.2 ขั้นตอนการสกัดคุณลักษณะ

ในขั้นตอนการสกัดคุณลักษณะด้วยวิธี HOG ผู้วิจัยได้ใช้ข้อมูลในการเก็บค่า ทิศทางค่าเกรเดียนต์ $0^\circ - 180^\circ$ โดยใช้ค่าทิศทางจำนวน 6 bin และกำหนดค่า กริด (grid) เป็นขนาด 8×8



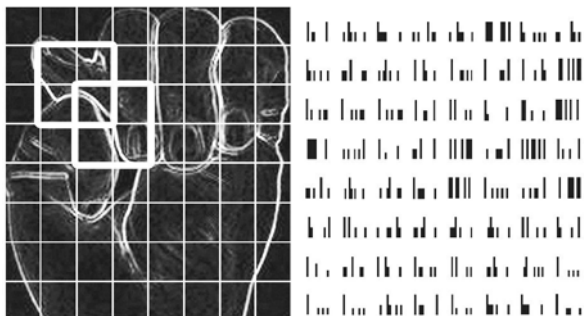
ภาพที่ 4: ตัวอย่างภาพท่ามือที่ใช้ในการทดลอง

เพื่อให้ได้ความแปรปรวนที่เหมาะสมผู้ทดลองได้ กำหนดให้ค่า v และ v' เป็นค่าปกติและไม่ปกติตามลำดับ โดยค่าสัมบูรณ์ v เป็นค่าแสดง L2 - norm ของเวกเตอร์ และ ϵ เป็นค่าคงที่ป้องกันข้อผิดพลาดที่เกิดจากการหารด้วยศูนย์ ดังสมการ

$$v' = v / (\|v\| + \epsilon) \tag{10}$$

และนอร์มอลไลซ์ด้วยการคำนวณตัวชี้วัดค่าความเข้มจากโอเวอร์แลปของเซลล์ภายในบล็อก ทำการยูเนียนเซลล์ขนาด 2×2 เพื่อใช้ในการลดผลกระทบจากการเปลี่ยนแปลงของแสงและเงาให้น้อยลง ค่าฮิสโตแกรมที่ถูกรับทิศทางค่าเกรเดียนท์ไว้จะ

ใช้เป็นตัวแทนแสดงคุณลักษณะในแต่ละเซลล์ และเมื่อเสร็จสิ้นกระบวนการ การเปลี่ยนรูป (Transformation) จะได้มิติของข้อมูล (Dimension) ขนาด 384 มิติซึ่งได้มาจากจำนวนเซลล์ย่อยคูณกับค่าทิศทาง [8]



ภาพที่ 5: ภาพตัวอย่างการโอเวอร์แลปบล็อกและการออร์มอลไลซ์ค่าความถี่ของทิศทางตามค่าเกรเดียนท์

หลังจากนั้นนำข้อมูลที่ได้นำเข้าสู่กระบวนการลดข้อมูลที่ไม่จำเป็นออกด้วยการวิเคราะห์องค์ประกอบหลัก โดยใช้คำนวณหาค่า Covariance matrix แล้วทำการวิเคราะห์แยกส่วนค่าไอเกน ด้วยวิธี Singular Value Decomposition (SVD) โดยเป็นการแยกเมตริก ออกเป็นผลคูณ 3 เมตริกด้วยสมการ

$$A = QSV \tag{11}$$

โดย Q และ V เป็น orthonormal matrix และ S เป็น diagonal matrix โดยใช้คุณสมบัติสำคัญของการแยกส่วนหาค่าสัมประสิทธิ์ที่เหมาะสม จากนั้นนำค่า Eigenvalues ที่ได้มาทำการบวกเพิ่มขึ้นเรื่อยๆ จนกระทั่งได้ค่าที่มากกว่าหรือเท่ากับค่า threshold ที่กำหนดไว้ ซึ่งจะทำให้ระบบสามารถรู้จำได้แม่นยำยิ่งขึ้น เมื่อจากผ่านกระบวนการวิเคราะห์องค์ประกอบหลักแล้วจะสามารถลดมิติของข้อมูลลงเหลือ 181 มิติ เพื่อหาประสิทธิภาพของการสกัดคุณลักษณะที่ผู้วิจัยได้ทดสอบการรู้จำเบื้องต้นด้วยวิธีหาค่าความเหมือนของข้อมูล ด้วยวิธีสหสัมพันธ์ไขว้ปรับปกติ (Normalized Cross Correlation: NCC) โดยมีอัตราค่าฝึกสอนและทดสอบเท่ากับ 3 : 2 และการหาประสิทธิภาพอัตราค่าความถูกต้องจากสมการ

$$AC_{RATE} = \frac{M_C}{M_T} \times 100 \tag{12}$$

เมื่อ AC_{RATE} คืออัตราค่าความถูกต้อง M_C คือข้อมูลที่รู้จำได้ถูกต้อง และ M_T คือจำนวนภาพทั้งหมด ผลการวัดประสิทธิภาพที่ได้แสดงใน ตารางที่ 1

ตารางที่ 1: ตารางเปรียบเทียบประสิทธิภาพในขั้นตอนการสกัดคุณลักษณะ

Method	PCA	HOG	HOG - PCA
Dimension	373	384	181
Accuracy	68 %	90 %	91 %

3.3 ขั้นตอนการจำแนกข้อมูล

ในการจำแนกข้อมูลด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีน แบบเรดิยัลเบสิสฟังก์ชันเคอร์เนล ผู้วิจัยได้ทำการจัดเตรียมข้อมูลโดยกำหนด อัตราค่าฝึกสอนและทดสอบเท่ากับ 3 : 2 เพื่อทดสอบเปรียบเทียบประสิทธิภาพของผลการรู้จำและประสิทธิภาพด้านระยะเวลาในการประมวลผลระหว่างขั้นตอนวิธีการ PCA - SVM, HOG - SVM และ HOG - PCA - SVM

4. สรุปผลการทดลองและอภิปรายผล

ผลที่ได้จากการทดลองปรากฏว่าขั้นตอนวิธีการ HOG - PCA - SVM มีประสิทธิภาพการรู้จำสูงสุดโดยมีความถูกต้องในการรู้จำเท่ากับร้อยละ 96 และมีผลรวมระยะเวลาที่ใช้ในการจำแนกที่น้อยสุดคือ 59.84 วินาที แสดงผลดังตารางที่ 2

ตารางที่ 2: ตารางผลการทดลอง

Method	PCA - SVM	HOG - SVM	HOG - PCA - SVM
Attributes	373	384	181
Times (s.)	69.07	72.18	59.84
Accuracy	72 %	95 %	96 %

โดยขั้นตอนวิธีการที่ผู้วิจัยนำเสนอสามารถลดขนาดของข้อมูล คิดเป็นร้อยละ 51.47 สามารถลดระยะเวลาในการประมวลผลคิดเป็นร้อยละ 13.36 เมื่อเปรียบเทียบกับขั้นตอนวิธีการ PCA - SVM และสามารถลดขนาดของข้อมูล คิดเป็นร้อยละ 52.86 สามารถลดระยะเวลาในการประมวลผล คิด

เป็นร้อยละ 17.10 เมื่อเปรียบเทียบกับขั้นตอนวิธีการ HOG - SVM

5. ข้อเสนอแนะ

เพื่อให้การจำแนกข้อมูลด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนแบบเรเดียลเบสิสฟังก์ชันเคอร์เนล เกิดประสิทธิภาพสูงสุดควรมีวิธีการที่เหมาะสมในการปรับค่าพารามิเตอร์

เอกสารอ้างอิง

- [1] ม่านฟ้า สุวรรณรัตน์, วิไลภรณ์ รุ่งศรีทอง, อนุชา รัตนสินธุ, ลอย แอนเดอร์สัน และโอเวน ที. ริก ลีย์ (บรรณาธิการ) “*ปทานุกรมภาษามือไทย ฉบับปรับปรุงและขยายเพิ่มเติม*” สมาคมคนหูหนวกแห่งประเทศไทย, กรุงเทพมหานคร พ.ศ. 2533.
- [2] M. Ebrahim Al-Ahdal and Nooritawati Md Tahir, “Review in Sign Language Recognition Systems Computers & Informatics (ISCI)” *IEEE Symposium*, pp.52-57, 2012.
- [3] N. Ditcharoen, K. Naruedomkul, N. Cercone “SignMT: An alternative language learning tool,” *Published in Journal Computers & Education*, vol.55, Issue 1, pp. 118-130, August 2010.
- [4] N. Ditcharoen, N. Cercone, K. Naruedomkul, and B. Tipakorn, “TSTMT: Step towards an accurate thai sign translation,” *Proceedings of the IEEE Fourth International Conference on Machine Learning and Applications(ICMLA)*, 2005.
- [5] Deng-Yuan Huang, Wu-Chih Hu, Sung-Hsiang Chang “Vision-Based Hand Gesture Recognition Using PCA+Gabor Filters and SVM,” *Intelligent Information Hiding and Multimedia Signal Processing, Proceedings of the IEEE International Conference*, pp.1-4. ,2009.
- [6] Milad Moghaddam, Manoochehr Nahvi, Reza Hassanzadeh Pak, “Static Persian Sign Language Recognition using Kernel-based Feature Extraction,” *Machine Vision and Image Processing (MVIP), 7th Iranian*, pp. 1-5, 2011.
- [7] Joyeeta Singha and Karen Das, “Indian Sign Language Recognition Using Eigen Value Weighted Euclidean Distance Based Classification Technique,” *International Journal of Advanced Computer Science and Applications (IJASCSA)*, Vol. 4, No. 2, pp. 190-195, 2013.
- [8] Furkan Işıkdoğan and Songül Albayrak, “Automatic Recognition of Turkish Fingerspelling,” *Innovations in Intelligent Systems and Applications (INISTA), International Symposium*, pp. 264-267, 2011.
- [9] Karishma Dixit and Anand Singh Jalal, “Automatic Indian Sign Language Recognition System,” *3rd IEEE International Advance Computing Conference(IACC)*, pp. 883-887, 2013.
- [10] Supawadee Saengsri, Vit Niennatrakul and Chotirat Ann Ratanamahatana, “TFRS: Thai finger-spelling sign language recognition system,” *Second International Conference, Digital Information and Communication Technology and it's Applications (DICTAP)*, pp. 457- 462, 2012.
- [11] Yahia Said, Mohamed Atri and Rached Tourki, “Human Detection Based on Integral Histograms of Oriented Gradients and SVM,” *International Conference, Communications, Computing and Control Applications (CCCA)*, pp. 1- 5, 2011.
- [12] Chumpol Senapan and Khamron Sunut, “Human Detection using Census Transform-Histogram of Oriented Gradients based feature and a Modified Version of the Extreme Learning Machine for Generalized Radial Basis Function Neural Networks,” *The 9th National Conference on Computing and Information Technology (NCCIT)*, pp. 677- 682, 2013.
- [13] นริศรา ดอนพรพันธ์, สิริภัทร เชี่ยวชาญวัฒนา และ กำรณ สุนันติ, “การเปรียบเทียบประสิทธิภาพการหาขอบภาพเพื่อการสกัดลักษณะเด่นด้วย PHOG สำหรับการจำแนกประเภทข้อมูลรูปภาพ” *The 7th International Joint Conference on Computer Science and Software Engineering Communications (JCSSE)*, 2010.
- [14] Sompong Valuvanathorn, Supot Nitsuwat and Mao Lin Huang “Multi-Feature Face Recognition based on PSO -SVM,” *IEEE conference publication, International Conference, ICT and Knowledge Engineering (ICT & Knowledge Engineering)*, pp. 140- 145, 2012.

เออาร์ทรอนิกส์: การรู้จำวัตถุของเทคโนโลยีความเป็นจริงเสริมโดยใช้อัลกอริทึม SURF สำหรับการประกอบอุปกรณ์ลงแผงวงจรพิมพ์ ARtronics: An Object-based Recognition of Augmented Reality for PCB Assembly using SURF Algorithm

กฤตชัย บุญศิวนนท์ (Krittachai Boonsivanon)¹ และอานุภาพ มีสมบุญ (Anupap Meesomboon)²

^{1,2} ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ มหาวิทยาลัยขอนแก่น

¹krittachai@fci.ksu.ac.th, ²anupap@kku.ac.th

บทคัดย่อ

การพัฒนาโปรแกรมด้วยเทคโนโลยีความเป็นจริงเสริมอาศัยวิธีการรู้จำรูปแบบด้วยเครื่องหมายเออาร์ (AR marker) ซึ่งพบปัญหาและข้อจำกัดในการใช้งานเป็นอย่างมาก โดยเฉพาะการนำมาใช้ในการพัฒนาร่วมกับการประกอบอุปกรณ์ลงแผงวงจรพิมพ์ ซึ่งมีขนาดเล็ก งานวิจัยนี้จึงนำเสนอการรู้จำวัตถุของเทคโนโลยีความเป็นจริงเสริมโดยประยุกต์ใช้อัลกอริทึม SURF สำหรับการประกอบอุปกรณ์ลงแผงวงจรพิมพ์อิเล็กทรอนิกส์ เรียกว่า “เออาร์ทรอนิกส์” และวัดประสิทธิภาพความถูกต้องของการรู้จำ โดยการทดลองการรู้จำใช้ชุดข้อมูลทดสอบเป็นรูปภาพแบบจำลองแผงวงจรพิมพ์ขนาด 640x480 จุดภาพ จำนวน 5 ชุด โดยในแต่ละชุดจะมีรูปภาพของวัตถุในภาพมีมุมมอง และการพรั่มัวของภาพที่แตกต่างกัน จำนวนชุดละ 10 ภาพ และใช้การวัดประสิทธิภาพของการรู้จำ โดยการหาค่าความถูกต้อง (Accuracy) และค่าระลึก (Recall) ผลการทดลองพบว่าเออาร์ทรอนิกส์มีค่าเฉลี่ยความถูกต้องและค่าการระลึกของชุดทดสอบทั้งหมด เท่ากับ 84.13% และ 87.0% ตามลำดับ ดังนั้นเออาร์ทรอนิกส์จึงมีประสิทธิภาพในการรู้จำวัตถุ และการประมวลผลด้วยเทคโนโลยีความเป็นจริงเสริมได้อย่างถูกต้อง

คำสำคัญ: อัลกอริทึม SURF การรู้จำวัตถุ ความเป็นจริงเสริม การประกอบอุปกรณ์ลงแผงวงจร การสกัดคุณลักษณะเด่นและจับคู่ภาพ

Abstract

Many Augmented Reality (AR) applications have utilized pattern recognition techniques. AR marker-based recognition leads to many problems and limitation, especially in Printed Circuit Board (PCB) assembly of

small size. In this paper, we propose an object-based recognition of AR for PCB assembly using Speeded-Up Robust Feature (SURF) algorithm called “ARtronics” and compare the accuracy and recall rate of five data sets. Each data set consists of ten 640x480 images with different viewpoints and blurriness. The results show that the average values of accuracy and recall are 84.13% and 87.0%, respectively. Hence, the ARtronics looks promising to be used for recognizing the PCB assembly.

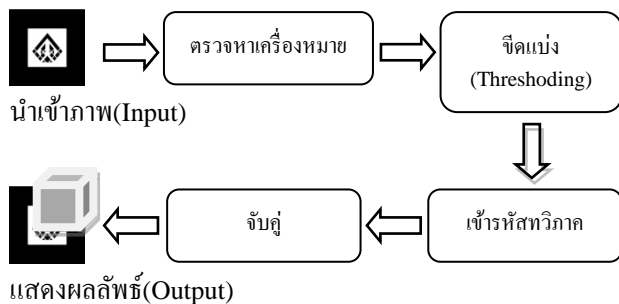
Keywords: SURF algorithm, Object-based recognition, Augmented Reality, PCB Assembly, Feature detection and matching.

1. บทนำ

การประกอบอุปกรณ์ลงบนแผงวงจรพิมพ์อิเล็กทรอนิกส์ (Printed Circuit Board Assembly; PCBA) จำเป็นต้องอาศัยทักษะและความเข้าใจเกี่ยวกับชิ้นส่วนอุปกรณ์อิเล็กทรอนิกส์ โดยเฉพาะปัญหาด้านเทคนิคการประกอบแผ่นวงจรพิมพ์ ซึ่งหากไม่ได้รับความฝึกฝนหรือมีทักษะในด้านดังกล่าว อาจส่งผลให้อุปกรณ์หรือแผ่นวงจรพิมพ์เสียหายได้ นอกจากนี้ยังพบว่าปัญหาที่เกี่ยวกับ โปรแกรมหรือสื่อที่ช่วยในการเรียนรู้ที่สามารถช่วยสอนด้วยการปฏิบัติการจริง ยังคงต้องอาศัยการดูจากรูปภาพ ซึ่งทำให้ไม่ได้รับความสะดวกและยากในการเข้าใจ จึงอาจทำให้เกิดความเบื่อหน่ายไม่ตั้งใจมากนัก อาจส่งผลให้ผู้เรียนไม่สนใจหรือเข้าใจในเนื้อหาเท่าที่ควร จึงจำเป็นต้องอาศัยเทคโนโลยีใหม่ที่เหมาะสมสำหรับใช้ในการแก้ปัญหาข้างต้น

เทคโนโลยีเออาร์หรือความเป็นจริงเสริม (Augmented Reality; AR) [1-3] เป็นเทคโนโลยีที่มีความสามารถในการแสดงกราฟิกจำลองเสมือนจริง 3 มิติอาศัยเทคนิคการแสดงผลทัศนศาสตร์คอมพิวเตอร์ (Computer vision) ร่วมกับเทคนิค

การประมวลผลภาพดิจิทัลที่ใช้ในการติดตาม (Tracking) และการรู้จำรูปแบบ (Pattern recognition) ด้วยเครื่องหมายเออาร์ (AR marker) [2-3] เพื่อเข้ารหัสทวิภาค(Binary code) ด้วยวิธีขีดแบ่ง(Thresholding)เพื่อใช้ในการคำนวณตำแหน่งระหว่างกล้องกับเครื่องหมายเออาร์ รวมทั้งเพื่อระบุและนำคุณลักษณะจากการเข้ารหัส (Identification marker; IDs) นำมาจับคู่กับเทมเพลตในหน่วยความจำคอมพิวเตอร์หรือฐานข้อมูลในรูปแบบการส่งข้อมูลสัญญาณวิดีโอต่อเนื่องชนิดการถ่ายทอดทันทีกาล (Live-video streams) ดังภาพที่ 1



ภาพที่ 1: กระบวนการรู้จำรูปแบบด้วยเครื่องหมาย [2]

อย่างไรก็ตาม ปัญหาของวิธีการรู้จำรูปแบบด้วยเครื่องหมายของเทคโนโลยีดังกล่าว คือไม่สะดวกในการใช้งานและการติดตั้งเครื่องหมาย และยังมีข้อจำกัดอื่นๆ ที่เป็นสาเหตุหลัก และส่งผลให้เกิดความผิดพลาดในการติดตามและรู้จำ เช่น ขนาดของเครื่องหมาย มุม และระยะห่างที่ใช้ในการติดตาม การปิดบังการติดตาม และความส่องสว่าง [3-4] ซึ่งส่งผลให้ไม่สามารถรู้จำและประมวลผลได้

จากปัญหาข้างต้น จึงไม่สามารถนำเทคโนโลยีดังกล่าวมาประยุกต์ใช้กับแผงวงจรพิมพ์ เนื่องจากแผงวงจรพิมพ์มีขนาดเล็ก ดังนั้นงานวิจัยนี้จึงได้นำเสนอการพัฒนาการรู้จำของเทคโนโลยีความเป็นจริงเสริมโดยใช้อัลกอริทึม SURF สำหรับการประกอบอุปกรณ์ลงแผงวงจรพิมพ์อิเล็กทรอนิกส์ด้วยการรู้จำพื้นฐานวัตถุ เรียกว่า “เออาร์ทรอนิกส์” และวัดประสิทธิภาพความถูกต้องของการรู้จำ ส่งผลให้ได้อัลกอริทึมที่สามารถใช้ในการรู้จำวัตถุในภาพแทนการรู้จำรูปแบบเครื่องหมาย

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 เทคโนโลยีความเป็นจริงเสริม

เทคโนโลยีความเป็นจริงเสริม มีการใช้งานร่วมกับอุปกรณ์แสดงผลอื่นๆ เช่น Head Mount Display (HMDs) อุปกรณ์ชนิดพกพา จอภาพแสดงผล และสมาร์ทโฟน เป็นต้น ปัจจุบันการรู้จำสามารถจำแนกออกเป็น 2 ประเภท คือ วิธีการติดตามและรู้จำด้วยเครื่องหมาย (Marker-based) และวิธีการการติดตามและรู้จำโดยไม่มีเครื่องหมาย (Markerless-based) ได้แก่ Image-based, Sensor-based, Location-based, Video-based และ Object-based เป็นต้น โดยวิธีการที่ได้รับความนิยม คือ วิธีการติดตามและรู้จำพื้นฐานเครื่องหมายเออาร์ เนื่องจากมีข้อดี คือ ค่าใช้จ่ายไม่สูงและง่ายต่อการติดตั้งและทดสอบ และมีไลบรารีหรือชุดซอฟต์แวร์ช่วยพัฒนาที่ใช้ในการพัฒนาโปรแกรมอยู่เป็นจำนวนมากในปัจจุบันทั้งที่เป็นโปรแกรมรหัสเปิด (Open source) และ โปรแกรมลิขสิทธิ์ เช่น ไลบรารี ARToolkit [2], ARTag [5], Metaio SDK [6] เป็นต้น

2.2 การรู้จำด้วยอัลกอริทึม SURF

การรู้จำ (Recognition) [7] เป็นหัวข้อหนึ่งทางด้านทัศนศาสตร์คอมพิวเตอร์ที่ทำหน้าสอนให้เครื่องเรียนรู้ (Machine learning) เกี่ยวกับการจับคู่วัตถุจากภาพสองภาพที่มีวัตถุในภาพเหมือนกัน โดยมีขั้นตอนหลักๆ ได้แก่ การสกัดคุณลักษณะเด่น (Feature extraction) และการจับคู่ภาพ (Key point matching) สำหรับอัลกอริทึมที่ได้รับความนิยมในการศึกษาเพื่อพัฒนาปรับปรุงประสิทธิภาพในการรู้จำหรือจำแนกวัตถุจากภาพ ได้แก่ SIFT [8] , PCA-SIFT [9] , GSIFT [9] และ SURF เป็นต้น

อัลกอริทึม Speeded-Up Robust Feature (SURF) [10-11,13] เป็นอัลกอริทึมหนึ่งที่มีประสิทธิภาพในการรู้จำได้รวดเร็ว โดยอาศัยการคำนวณเพื่อตรวจหาจุดสนใจ (Key point) ในภาพที่มีการเปลี่ยนแปลงเชิง 2 มิติของระดับความเข้มแสงรอบๆจุดสนใจนั้น ซึ่งอัลกอริทึมนี้ได้รับความนิยมและมีผู้นนำมาศึกษาและพัฒนาปรับปรุงประสิทธิภาพอย่างต่อเนื่อง โดยเฉพาะในการพัฒนาในหัวข้อของการรู้จำวัตถุจากภาพ (Object-based recognition) โดยมีขั้นตอนการทำงานหลักๆ 3 ขั้นตอน ได้แก่ การตรวจหาจุดสนใจ (Key point detection) การอธิบายจุดสนใจ (Key point description) และการจับคู่ภาพ (Matching)

ขั้นตอนหนึ่งที่สำคัญสำหรับการรู้จำวัตถุจากภาพ คือ ขั้นตอนการสกัดคุณลักษณะจากภาพทั้ง 2 ภาพที่ต้องการ ด้วยวิธีการค้นหาจุดสนใจ และนำมาจับคู่กับภาพต้นฉบับ โดยอัลกอริทึมนี้คล้ายกับอัลกอริทึม SIFT เนื่องจากได้พัฒนาปรับปรุงประสิทธิภาพบนพื้นฐานทฤษฎีและขั้นตอนการปรับภาพพื้นฐานหรือการแปลงสัมพรรค (Affine transformation) ของ SIFT [9] ในขณะที่มีการปรับปรุงและเปลี่ยนแปลงวิธีการ และขั้นตอนการตรวจหา และการอธิบายจุดสนใจ จึงทำให้อัลกอริทึม SURF มีความทนทานต่อการเปลี่ยนแปลงที่เกิดขึ้นต่างๆ ได้ดีกว่าอัลกอริทึมอื่นๆ รวมถึง SIFT ด้วย เช่น ความเร็ว ความส่องสว่าง และการหมุน [9-10,13] เป็นต้น สาเหตุที่ SURF ทำงานได้เร็วกว่า เนื่องจากการประมวลผลที่ไม่ได้ค้นหาจุดสนใจบนภาพต้นฉบับ แต่หาจุดสนใจบนภาพอินทิกรัล (Integral image) [14] การคำนวณคอนโวลูทด้วยผลการคอนโวลูทกับตัวประมาณอนุพันธ์อันดับที่สองของเกาส์เซียน (Gaussian derivative) และการตรวจหาแบบ non-maximal-suppression ของดีเทอร์มิแนนต์ Hessian matrix ซึ่งนับได้ว่าเป็นจุดสำคัญที่สุดของอัลกอริทึมนี้ [11] ขั้นตอนการทำงานของอัลกอริทึม SURF มีทั้งหมด 4 ขั้นตอน ดังนี้

ขั้นตอนที่ 1 Key point detection เป็นขั้นตอนการตรวจหาจุดสนใจบนภาพอินทิกรัลเพื่อที่ใช้ในขั้นตอนถัดไป ดังสมการที่ 1

$$I\Sigma(P) = \sum_{i=0}^x \sum_{j=0}^y I(x, y) \tag{1}$$

อัลกอริทึม SURF ใช้การหาด้วยผลการคอนโวลูทกับตัวประมาณอนุพันธ์อันดับที่สองของเกาส์เซียนของภาพต้นฉบับบนปริภูมิ Laplacian of Gaussian (LoG) หรือเรียกว่า Hessian matrix ดังสมการที่ 2

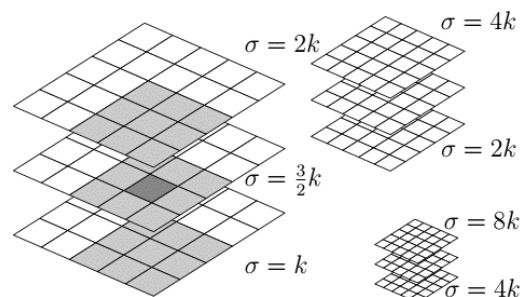
$$H(x, \sigma) = \begin{bmatrix} Lxx(x, \sigma) & Lxy(x, \sigma) \\ Lxy(x, \sigma) & Lyy(x, \sigma) \end{bmatrix} \tag{2}$$

โดยที่ $Lxx(x, \sigma)$, $Lyy(x, \sigma)$ และ $Lxy(x, \sigma)$ คือ ผลคูณการคอนโวลูทตัวประมาณอนุพันธ์อันดับที่สองของเกาส์เซียน ณ ตำแหน่ง x ในแนวแกน x , y และ $x-y$ ตามลำดับ จากนั้นกำหนดจุดภาพใดๆที่ใช้เป็นจุดสนใจ จึงทำให้คำนวณได้เร็ว ดังสมการที่ 3

$$\det(H_{\text{approx}}) = D_{xx} D_{yy} - (0.9D_{xy})^2 \tag{3}$$

โดยที่ D_{xx} , D_{yy} และ D_{xy} เป็นการประมาณอนุพันธ์อันดับที่ 2 ของเกาส์เซียนในทิศทาง คือ x , y และ $x-y$ ตามลำดับ เพื่อให้จุดภาพสี่เทาบริเวณรอบจุดสนใจมีค่าเท่ากับ 0 ส่งผลให้การคำนวณมีประสิทธิภาพมากขึ้น และค่า 0.9 คือ ค่าที่ใช้สำหรับปรับสมมติเดเทอร์มิแนนต์ของ Hessian matrix สำหรับตัวกรองแบบ Box filters และภาพอินทิกรัลมาคอนโวลูทซ้ำในชั้นก่อนหน้า แล้วจึงนำมาคอนโวลูทบนภาพต้นฉบับ นอกจากนี้ยังสามารถปรับสเกล (σ) ของแก่นกลาง Gaussian kernel และขนาดของตัวกรองแบบ Box filters ได้ด้วย Logarithmic scale

ขั้นตอนที่ 2 Key point Localization เป็นขั้นตอนในการหาจุดสนใจ โดยนำค่าที่ได้กำหนดจุดสนใจ 1 จุดภาพ จากนั้นอาศัยการคำนวณค่าของแต่ละจุดภาพรอบๆที่ใกล้เคียงที่สุด (Nearest neighbor) ที่อยู่โดยรอบจุดสนใจ มาทำการคำนวณหาค่าสูงสุด (Local maxima) หรือค่าต่ำสุด (Local minima) ของแต่ละจุดภาพบนปริภูมิ LoG โดยแบ่งเป็น 3 ระดับ จากนั้นจุดสูงสุดหรือจุดต่ำสุดที่ได้จะถูกนำมาจับคู่กับจุดภาพรอบๆอีก 8 จุดภาพใกล้เคียงที่สุดกับจุดสนใจที่ได้กำหนดในข้างต้นในระดับเดียวกัน และทำการคำนวณขนาด 2×9 ส่วน (18 จุดภาพ) ที่อยู่รอบๆบนปริภูมิ LoG ในสเกลใกล้เคียงในอีก 2 ระดับที่เหลือ ซึ่งวิธีการดังกล่าว คือ การตรวจหาแบบ non-maximal-suppression ขนาด $3 \times 3 \times 3$ ส่วน ดังภาพที่ 2 ซึ่งหากว่าจุดภาพที่ได้เป็นจุดสูงสุด หรือจุดต่ำสุด ก็จะถูกเลือกเป็นจุดสนใจแบบ key point candidate เพื่อใช้ในการคำนวณหาจุดสนใจที่มีแนวโน้มจะเป็นขอบภาพในขั้นตอนถัดไป



ภาพที่ 2: การหาจุดสนใจ (Key point localization) [11]

ขั้นตอนที่ 3 Orientation Assignment เป็นขั้นตอนที่ใช้คุณสมบัติของ Harr Wavelets ด้วยเวกเตอร์ในแต่ละกลุ่มย่อย

และการกวาดไปรอบๆเป็นวงกลมทุกๆค่าที่ใกล้เคียงที่สุดกับจุดกลุ่มตัวอย่างได้จากกรคำนวณด้วย Harr Wavelets แล้วจึงหาค่าผลรวมเวกเตอร์ของจุดสนใจ และใช้เป็น Dominate Orientation โดยอาศัยการคำนวณระยะทางระหว่างตัวตรวจหากับจุดสนใจด้วยการวัดระยะทางแบบยูคลิด (Euclidean distance)

ขั้นตอนที่ 4 *Generation of key point descriptors* เป็นขั้นตอนสำหรับแบ่งบริเวณรอบจุดสนใจออกเป็นขนาด 3x3 ส่วน โดยแต่ละส่วนจะทำการคำนวณหาค่าผลรวมเวกเตอร์ที่สูงสุดในทิศทาง x และ y จากนั้นทำการให้ค่าน้ำหนักค่าผลรวมแก่นกกลาง Gaussian kernel โดยให้อยู่ตำแหน่งกลางบริเวณจุดสนใจ จากนั้นจึงคำนวณหาค่าผลรวมในทิศทาง x และ y และค่าสัมประสิทธิ์ในทิศทาง x และ y จึงจะได้เวกเตอร์ตัวอธิบาย (Descriptors vector) ที่สามารถแบ่งออกเป็นขนาด 3x3x3 ส่วน โดยสามารถคำนวณได้จากค่าความเข้มแสงของจุดภาพในบริเวณใกล้เคียงที่สุรอบๆจุดสนใจ ซึ่งเวกเตอร์อธิบายที่ได้จะถูกนำมาใช้ในขั้นตอนการจับคู่และรู้จำอัตลักษณ์ (Identity) กับภาพต้นฉบับหรือภาพอื่นๆต่อไป

2.3 งานวิจัยที่เกี่ยวข้อง

การโต้ตอบบนระบบ e-Learning รายวิชา ภาษาอังกฤษ และวิทยาศาสตร์ โดยอาศัยการรู้จำรูปแบบและความเป็นจริงเสริม [15] โดยใช้เทคนิคการรู้จำรูปภาพด้วยอัลกอริทึม SURF และเนื้อหาที่เป็นความเป็นจริงเสริม กราฟิกเคลื่อนไหวหรือภาพยนตร์ สำหรับการโต้ตอบบนระบบ e-Learning ใช้เครื่องหมาย color-band หรือ polka-dot แทนตำแหน่งเคอร์เซอร์ของเมาส์โดยใช้นิ้วมือในการชี้เพื่อระบุตำแหน่งในหนังสือรูปภาพที่ใช้ในการเรียนรู้ และใช้ระบุตำแหน่งที่จะใช้สร้างวัตถุเสมือนจริง ผลการทดลองพบว่า ระบบนี้ได้รับความนิยมและสนใจในการเรียนรู้มากขึ้น

ระบบช่วยในการวางแผนและการออกแบบการประกอบผลิตภัณฑ์ด้วยเทคโนโลยีความเป็นจริงเสริม [16] ได้นำเสนอด้วยการใช้ขั้นตอนการวางแผนในการออกแบบผลิตภัณฑ์ และขั้นตอนการวางแผนพื้นที่สำหรับใช้ในการออกแบบ โดยข้อมูลของขั้นตอนในการออกแบบ สามารถส่งกลับมายังผู้ออกแบบได้ทันที รวมทั้งช่วยให้ผู้ออกแบบสามารถแก้ไข และใส่ค่าที่ต้องการได้โดยตรงด้วยคีย์บอร์ดเสมือนจริง

ระบบตรวจหาป้ายจราจร โดยใช้อัลกอริทึม SURF แบบทันกาลบนบอร์ด FPGA ตระกูล Kintex-7 [17] โดยทำการออกแบบในส่วนของฮาร์ดแวร์ที่มีกระบวนการวัดความแม่นยำและพัฒนาส่วนของอัลกอริทึมด้วย SURF ในการตรวจหาป้ายจราจร โดยทำการทดลองความทนทานต่อการเปลี่ยนแปลงที่เกิดขึ้นต่างๆ ได้แก่ การหมุนภาพ และความส่องสว่าง ผลการทดลองพบว่า ระบบสามารถตรวจหาป้ายจราจรได้แม่นยำในรูปแบบสัญญาณวิดีโอ 800x600 ที่อัตรา 60 เฟรมต่อวินาที

การศึกษาและเปรียบเทียบประสิทธิภาพของอัลกอริทึมที่พัฒนาปรับปรุงอัลกอริทึม SIFT [9] ได้แก่ PCA-SIFT, GSIFT, CSIFT, ASIFT และ SURF และวัดอัตราความถูกต้องในการจับคู่อัลกอริทึมทั้งหมด โดยใช้ชุดภาพทดสอบ (Dataset) ของ Visual Geometry Group และทดสอบความทนทานต่อการเปลี่ยนแปลงที่เกิดขึ้นต่างๆ ได้แก่ ขนาดและการหมุนภาพ การเบลอภาพ การแปลงสัมพรรค และความส่องสว่าง ผลการทดลองพบว่า เวลาเฉลี่ยที่ใช้ในการสกัดคุณลักษณะและจับคู่ภาพ อัลกอริทึม SURF ทำงาน ระดับดีเยี่ยม ในขณะที่การทำงานอื่นๆ ในระดับพอใช้

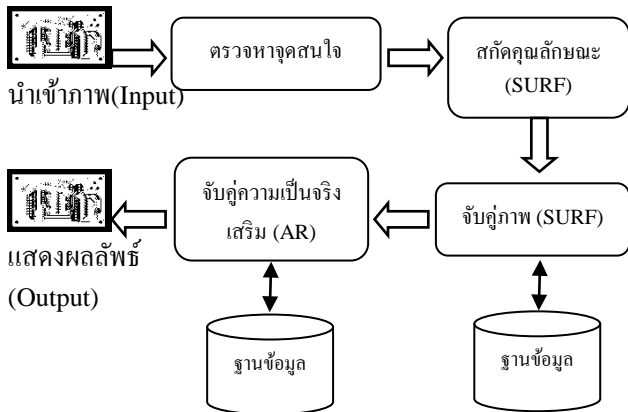
การศึกษาและเปรียบเทียบอัลกอริทึม SIFT, PCA-SIFT และ SURF [13] โดยใช้วิธีการ K-NN และ RANSAC ในการวิเคราะห์ผลการรู้จำ ในการทดลองความทนทานต่อการได้แก่ ขนาดภาพ การหมุนภาพ การเบลอภาพ และความส่องสว่าง โดยใช้วิธีการวัดค่าประสิทธิภาพ (Repeatability) และเวลาที่ใช้ ผลการทดลองพบว่า SURF ทำงานได้เร็วที่สุด และมีความทนทานต่อการเปลี่ยนแปลงที่เกิดขึ้นต่างๆ เช่น การหมุน และความส่องสว่างเมื่อเทียบกับอัลกอริทึมอื่นๆที่ใช้ในการทดลอง

จากการทบทวนงานวิจัยที่เกี่ยวข้อง พบว่างานวิจัยส่วนใหญ่พัฒนาโปรแกรมด้วยเทคโนโลยีความเป็นจริงเสริม โดยใช้การรู้จำรูปแบบด้วยเครื่องหมาย ยังไม่พบว่ามีการนำเอาอัลกอริทึม SURF ไปประยุกต์ใช้ในการพัฒนา นอกจากนี้ยังพบว่าอัลกอริทึมดังกล่าวมีประสิทธิภาพในการตรวจหา และจับคู่ภาพได้อย่างรวดเร็ว [9,13] จึงมีความเหมาะสมกับงานวิจัยนี้ เนื่องจากการนำไปประยุกต์ใช้ในรูปแบบสัญญาณวิดีโอต่อเนื่อง จึงจำเป็นต้องใช้อัลกอริทึมที่สามารถรู้จำได้รวดเร็ว

3. วิธีการดำเนินการวิจัย

3.1 การออกแบบ “เออาร์เทอร์นิกส์”

การพัฒนา “เออาร์ทรอนิกส์” ด้วยอัลกอริทึม SURF ที่ได้
นำเสนอในบทความนี้มีโครงสร้างการพัฒนา ดังภาพที่ 3



ภาพที่ 3: โครงสร้างการพัฒนาเออาร์ทรอนิกส์

3.2 การเตรียมการทดลอง

การพัฒนา “เออาร์ทรอนิกส์” ด้วยอัลกอริทึม SURF เพื่อใช้
สำหรับทดลองการรู้จำ อาศัยการเขียนโปรแกรมด้วย Visual
C++ และไลบรารี VLFeat [18] และการพัฒนาอัลกอริทึมเพื่อ
ใช้สำหรับทดลองการรู้จำสำหรับเทคโนโลยีความเป็นจริงเสริม
ด้วย Metaio SDK[6] โดยทดลองบนเครื่องไมโครคอมพิวเตอร์
(ซีพียู Intel Pentium Dual-core 2.60 GHz, หน่วยความจำ 4
GB) ภายใต้ระบบปฏิบัติการวินโดวส์ 7 และกล้อง USB ความ
ละเอียด 3 ล้านพิกเซล

3.3 การทดลองการรู้จำ

การทดลองการรู้จำใช้ชุดข้อมูล (Dataset) เป็นรูปภาพ
แบบจำลองแผงวงจรพิมพ์ ขนาด 640x480 จุดภาพ จำนวน 5
ชุด โดยในแต่ละชุดจะมีรูปภาพของวัตถุในภาพที่มีมุมมอง และ
การรบกวนของภาพแตกต่างกัน จำนวนชุดละ 10 ภาพ โดยการ
หาค่าความถูกต้อง (Accuracy; A) และค่าระลึก (Recall; R)
จากกลุ่มภาพ จำนวน 2 ชุด ได้แก่ ชุดภาพใช้ฝึกสอน (Training
set) จำนวน 30 ภาพ และชุดภาพใช้ทดสอบ (Test set) จำนวน
20 ภาพ

3.4 การทดลองเทคโนโลยีความเป็นจริงเสริม

การทดลองการรู้จำสำหรับเทคโนโลยีความเป็นจริงเสริม
โดยใช้การประมวลผลกราฟิก (Rendering) ในอัตรา 30 เฟรม
ต่อวินาที และวัดความถูกต้องในการรู้จำต่อเนื่องที่จำนวน 300
เฟรม

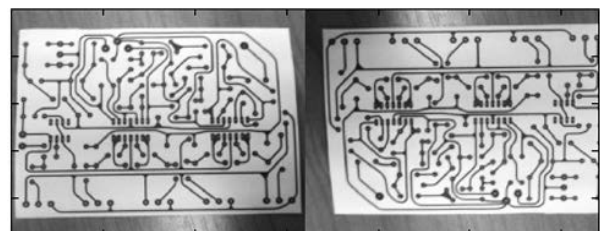
4. ผลการทดลอง

4.1 ผลการทดลองการรู้จำ

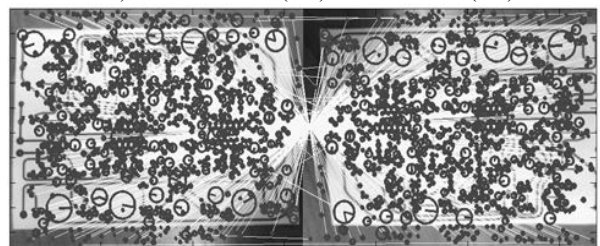
ผลการทดลองความถูกต้องและการระลึก พบว่า ชุดทดสอบ
ที่ 1 มีค่าความถูกต้องและค่าความระลึกดีที่สุด เท่ากับ 94.67%
และ 0.93 ตามลำดับ เนื่องจากวัตถุในภาพมีขนาดเล็ก และ
สวดลายน้อยจึงทำให้การรู้จำได้ดี และเออาร์ทรอนิกส์มีค่าเฉลี่ย
การทดลองชุดทดสอบทั้งหมด เท่ากับ 84.13% และ 0.87
ตามลำดับ ดังตารางที่ 1 และผลการประมวลผล ดังภาพที่ 4

ตารางที่ 1: ค่าเฉลี่ยของความถูกต้องและการระลึก

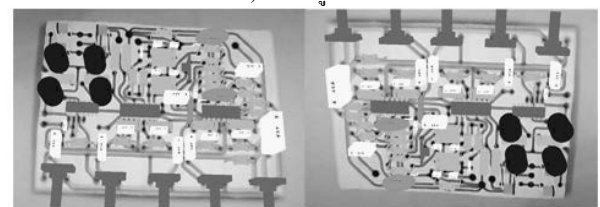
ชุดทดสอบ	ค่าความถูกต้อง (A)	ค่าระลึก (R)
ชุดทดสอบที่ 1	94.67%	0.93
ชุดทดสอบที่ 2	82.0%	0.86
ชุดทดสอบที่ 3	82.0%	0.86
ชุดทดสอบที่ 4	86.67%	0.88
ชุดทดสอบที่ 5	75.33%	0.81
ค่าเฉลี่ย	84.13%	0.87



ก) นำเข้าภาพต้นฉบับ (ซ้าย) และภาพทดสอบ (ขวา)



ข) การจับคู่ภาพ



ค) การประมวลผลกราฟิกด้วยเทคโนโลยีความเป็นจริงเสริม

ภาพที่ 4: ผลการทดลอง “เออาร์ทรอนิกส์”

5. สรุป

การพัฒนา “เออาร์ทรอนิกส์” โดยใช้อัลกอริทึม SURF โดยการทดลองการรู้จำใช้ชุดข้อมูลทดสอบเป็นรูปภาพแบบจำลองแผงวงจรพิมพ์ จำนวน 5 ชุด และหาค่าความถูกต้อง และค่าการระลึก ผลการทดลองพบว่า เออาร์ทรอนิกส์มีค่าเฉลี่ยความถูกต้องและค่าการระลึก เท่ากับ 84.13% และ 87.0% ตามลำดับ และสามารถประมวลผลกราฟิกด้วยเทคโนโลยีความเป็นจริงเสริมได้ถูกต้อง ส่งผลให้ได้อัลกอริทึมที่สามารถในการรู้จำด้วยวัตถุในภาพแทนการรู้จำรูปแบบเครื่องหมายเออาร์ นอกจากนี้จากการทดลองยังพบว่า ขนาด ลวดลายของวงจรและความส่องสว่างที่มีการเปลี่ยนแปลง ส่งผลให้การรู้จำได้ช้าและถูกต้องลดลง จึงควรพัฒนาปรับปรุงอัลกอริทึมนี้และเลือกใช้อัลกอริทึมอื่นๆ มาทำการทดลองภายใต้ความทนทานต่อการเปลี่ยนแปลงต่างๆ และเปรียบเทียบประสิทธิภาพต่อไป

เอกสารอ้างอิง

- [1] A. Van Dam, A. Forsberg, D. Laidlaw, J. LaViola and R. Simpson, “Immersive VR for Scientific Visualization,” *A Progress Report. IEEE Computer Graphics and Applications.*, Nov/Dec, 2000;Vol.20, No.6, pp.26-52.
- [2] H. Kato, M. Billinghurst and I. Poupyrev, “Virtual Object Manipulation on a Table-top AR Environment,” *Proceedings of the International Symposium on Augmented Reality(ISAR’2000)*, 2000, pp.111-119.
- [3] D. F. Abawi and J. Bienwald, “Accuracy in Optical Tracking with Fiducial Markers:An accuracy function for ARToolkit,” *Proceedings of the Third IEEE and ACM International Symposium on Mixed and Augmented Reality(ISMAR’2004)*, 2004, pp.260-261.
- [4] J. Seo, J. Shim, J. Choi, J. Park and T. Han, “Enhancing Marker-based AR Technology,” *Visual and Mixed Reality, Part I, HCII(2011)*, 2011, pp.97-104.
- [5] M. Fiala, “The ARTag Augmented Reality System Rev2h,” *Institute of Information Technology, NRC, Canada*, [serialonline]2006, March[cited 2009 Feb] Available from:URL:<http://www.artag.net>
- [6] Metaio, “Metaio SDK” *Metaio GmbH Inc.*, Munich, Germany, March[cited 2013 Jan] Available from:URL:<http://www.metaio.com/sdk/>
- [7] X. Zhang, Y. Yang, Z. Han, H. Wang and C. Gao, “Object Class Detection: A Survey,” *ACM Computing Survey*, Vol. 46, No. 1, pp. 1-54, 2013.
- [8] D. G. Lowe, “Distinctive Image Features From Scale-invariant Keypoints,” *International Journal of Computer Vision*, Vol. 60, No. 2, pp. 91-110, 2004.
- [9] J. Wu, Z. Cui, V. S. Sheng, P. Zhao, D. Su and S. Gong, “A Comparative Study of SIFT and its Variants,” *Measurement Science Review*, Vol. 13, No. 3, pp. 122-131, 2013.
- [10] H. Bay, T. Tuytelaars and L. V. Gool, “SURF:Speeded-Up Robust Features,” *Computer Vision—ECCV 2006: 9th European Conference on Computer Vision, 2006*, 7-13 May 2006, Spinker, Part II, pp.404-417.
- [11] J. T. Pederson, “Study group SURF:Feature Detection & Description,” *Department of Computer Science, Aarhus University, Aarhus, Denmark*, [cited 2013 Jan] Available from:URL: <http://cs.au.dk/~jtp/SURF/report.pdf>
- [12] R. Oji, “An Automatic Algorithm for Object Recognition and Detection Based on ASIFT Keypoints” *Signal & Image Processing: An International Journal(SIPIJ)*, Vol. 3, No. 5, pp. 29-40, October 2009.
- [13] L. Juan, O. Gwun, “A Comparison of SIFT, PCA-SIFT and SURF,” *International Journal of Image Processing*, Vol. 3, No. 4, pp. 143-152, 2009.
- [14] P. Viola and M. Jones, “Rapid Object Detection using a Boosted Cascade of Simple Features,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR’11)*, 2001.
- [15] S. H. Lee, J. Choi, and J. Park, “Interactive e-Learning System using Pattern Recognition and Augmented Reality,” *IEEE Transactions on Consumer Electronic*, Vol. 55, No. 2, pp. 883-890, May 2009.
- [16] Y. Ping et al., “Augmented Reality Aided Assembly Design and Planning,” *Annals of CIRP*, Vol. 56, pp. 49-52, 2007.
- [17] J. Zhao, S. Zhu, X. Huang, “Real-time Traffic Sign Detection using SURF Features on FPGA,” *High Performance Extreme Computing Conference (HPEC), 2013 IEEE*, 10-12 Sept. 2013, pp.1-6.
- [18] A. Vedaldi and B. Fulkerson, “VLFeat-An Open and Portable Library of Computer Vision Algorithms,” *Proceedings of the 18th International Conference on Multimedia(MM’10)*, Firenze, Italy, October 25-29, 2010.

การจำแนกท่าทางขณะรับชมโทรทัศน์โดยใช้กล้องคิเนค

Postural Classification during Television Watching using Kinect

ธรรมศาสตร์ วิชาธรรม (Thammarsat Visutarom)¹ พรชัย มงคลนาม (Pornchai Mongkolnam)²

และ โจนathan โสอิน ชาน (Jonathan Chan)³

คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี

¹thammarsat@gmail.com, ²pornchai@sit.kmutt.ac.th, ³jonathan@sit.kmutt.ac.th

บทคัดย่อ

ภาวะถดถอยของผู้สูงอายุเป็นสาเหตุหนึ่งที่ทำให้ผู้สูงอายุต้องอยู่ติดบ้านมากยิ่งขึ้น แต่ในทางกลับกันนั้นปัจจุบันนี้บุคคลในครอบครัวมีความจำเป็นต้องออกไปทำงานนอกบ้านมากขึ้นตามไปด้วย จึงทำให้ผู้สูงอายุถูกทอดทิ้งให้อยู่เพียงลำพัง ซึ่งการศึกษาพฤติกรรมและกิจกรรมของผู้สูงอายุจึงกลายเป็นส่วนหนึ่งที่สังคมกำลังให้ความสนใจ การประยุกต์ใช้เทคโนโลยีเพื่อติดตามชีวิตประจำวันของผู้สูงอายุจึงเป็นอีกแนวทางในการช่วยให้บุคคลในครอบครัวสามารถเข้าใจถึงพฤติกรรมของผู้สูงอายุได้

บทความนี้จึงได้นำเสนอระบบการจำแนกท่าทางผู้สูงอายุในระหว่างการรับชมโทรทัศน์ซึ่งเป็นหนึ่งในกิจกรรมหลักของชีวิตประจำวันของผู้สูงอายุ โดยใช้กล้อง Kinect (คิเนค) สองตัวในการตรวจจับ และจำแนกท่าทางของผู้สูงอายุ โดยการจำแนกท่าทางสามารถช่วยเพิ่มรายละเอียดของการศึกษาพฤติกรรมและกิจกรรมของผู้สูงอายุให้แก่บุคคลในครอบครัวและแพทย์ประจำตัวผู้สูงอายุได้ และทำให้พวกเขาสามารถเข้าใจผู้สูงอายุได้ดียิ่งขึ้น

คำสำคัญ: การจำแนกท่าทาง กล้องคิเนค ผู้สูงอายุ รับชมโทรทัศน์ที่บ้าน

Abstract

As the ageing population increases globally, there are more and more elderly people who need to stay home due to physical and/or mental limitations. On the other hand, their family members need to work outside and may leave the elderly to stay alone at home. Knowing more about the behaviors and activities of the elderly has been one of the most sought-after studies in today's ageing society. Technology could be used to monitor and keep records of the daily lives of the elderly, especially while being at home, to help us better understand and innovate technologies to serve them.

We propose a postural classification system to be used to keep track of the elderly while watching television, which is one of most time-consuming activities at homes. Two Kinect cameras were used to classify various postures of the elderly, and the data can be used for future behavioral analyses. The postural classification and the subsequent behavioral analyses could help provide a better understanding of the elderly by the family members, physicians, and other health practitioners.

Keywords: Postural classification, Kinect Camera, Elderly, Home Television Watching

1. บทนำ

สังคมผู้สูงอายุ คือสังคมที่มีประชากรอายุ 60 ปีขึ้นไป คิดเป็นร้อยละ 10 จากจำนวนประชากรทั้งหมด ซึ่งประเทศไทยได้ก้าวเข้าสู่สังคมผู้สูงอายุตั้งแต่ พ.ศ. 2543 [1] แต่ปัจจุบันบุคคลในครอบครัวกลับมีความจำเป็นต้องออกไปทำงานนอกบ้าน แม้จำนวนประชากรผู้สูงอายุจะเพิ่มมากขึ้นเท่าใดการได้รับการดูแลก็ยังลดน้อยลง การติดตามพฤติกรรมของผู้สูงอายุจึงเป็นอีกแนวทางหนึ่งที่จะช่วยให้สามารถเข้าใจผู้สูงอายุได้มากยิ่งขึ้น ซึ่งเมื่อจำนวนประชากรผู้สูงอายุเพิ่มขึ้นความต้องการติดตามพฤติกรรมผู้สูงอายุก็มากขึ้นตามไปด้วย

ด้วยร่างกายที่เข้าสู่ภาวะถดถอยทำให้ผู้สูงอายุต้องอยู่ติดบ้าน กิจกรรมจึงถูกจำกัดอยู่เพียงแต่ในบ้าน การรับชมโทรทัศน์จึงกลายเป็นกิจกรรมที่น่าสนใจสำหรับผู้สูงอายุ [2, 3] ในงานวิจัยนี้สังเกตเห็นว่าการรับชมโทรทัศน์เป็นอีกกิจกรรมหนึ่งที่น่าติดตามซึ่งจะช่วยเพิ่มรายละเอียดในการติดตามพฤติกรรมของผู้สูงอายุ ซึ่งจะส่งผลให้บุคคลในครอบครัวสามารถเข้าใจในตัวผู้สูงอายุได้ดียิ่งขึ้น

การนำเทคโนโลยีเข้ามาประยุกต์เพื่อใช้ในการติดตามกิจกรรม และพฤติกรรมของผู้สูงอายุจึงเป็นเป้าหมายของงานวิจัย โดยการนำกล้อง Kinect มาใช้ในการจำแนกท่าทางระหว่างการชมโทรทัศน์ซึ่งจะช่วยให้เราสามารถเข้าถึงอากัปกริยาของผู้สูงอายุในระหว่างการรับชมโทรทัศน์ได้อย่างละเอียด นอกจากนี้การนำเทคโนโลยีเข้ามาประยุกต์ใช้ในชีวิตประจำวันของผู้สูงอายุเป็นอีกวิธีที่ช่วยยกระดับความเป็นอยู่ของผู้สูงอายุให้ดียิ่งขึ้น [4]

2. วรรณกรรมที่เกี่ยวข้อง

2.1 การติดตามพฤติกรรมของผู้สูงอายุ

การประยุกต์เทคโนโลยีเพื่อใช้ในการติดตามพฤติกรรมของผู้สูงอายุช่วยให้บุคคลภายในครอบครัวสามารถติดตามกิจกรรมประจำวันของผู้สูงอายุได้แม้ว่าจะไม่ได้อยู่ด้วยกัน โดยการติดตามพฤติกรรมอาจแบ่งออกได้เป็นหลายรูปแบบ เช่น การติดตามอุบัติเหตุที่อาจเกิดกับผู้สูงอายุ เช่น การหกล้ม [5, 6] การติดตามคุณภาพรวมของชีวิตประจำวันก็เป็นอีกหนึ่งวิธีการที่

ช่วยให้บุคคลในครอบครัวสามารถรับรู้ถึงกิจกรรมต่าง ๆ ในแต่ละวันได้ซึ่งมีหลายวิธีการในการนำเทคโนโลยีมาช่วยในการติดตามภาพรวมชีวิตประจำวัน เช่น การติดกล้องไว้ตามจุดต่างๆเพื่อติดตามกิจกรรมในแต่ละเวลาของผู้สูงอายุ [7] โดยในปัจจุบันมีหลายงานวิจัยที่เริ่มให้ความสำคัญกับความเป็นส่วนตัวในระหว่างการติดตามชีวิตประจำวันของผู้สูงอายุมากขึ้น เพื่อไม่ให้การติดตามชีวิตประจำวันของผู้สูงอายุกลายเป็นการทำลายความเป็นส่วนตัวของผู้สูงอายุลง [8, 9] ในงานวิจัยชิ้นนี้จะใช้กล้อง Kinect ซึ่งมีความสามารถในการตรวจจับท่าทางเพื่อใช้ในการจำแนกท่าทางของผู้สูงอายุโดยที่ไม่มีกรบันทึกภาพของผู้สูงอายุเพื่อให้คงไว้ซึ่งความเป็นส่วนตัวของผู้สูงอายุ

2.2 กล้อง Kinect







กล้อง Kinect เป็นอุปกรณ์ที่งานวิจัยนี้เลือกใช้ โดยกล้อง Kinect เป็นอุปกรณ์ในการบันทึกภาพเชิงลึก (Depth Image) จากทางบริษัท Microsoft โดยกล้อง Kinect มี Library Functions หลายตัวที่ช่วยในการแยกแยะร่างกายมนุษย์ออกจากพื้นหลังของภาพเชิงลึก และสามารถค้นหาจุดต่าง ๆ บนร่างกายเพื่อสร้างเป็นโครงร่างจำลอง (Skeleton) ไว้ใช้ในการตรวจจับการเคลื่อนไหวของมนุษย์ เช่น Windows SDK และ OpenNI เป็นต้น ซึ่งสามารถช่วยลดเวลาในการหาวิธีการแยกแยะร่างกายมนุษย์ออกจากพื้น และตรวจจับการเคลื่อนไหวได้ ด้วยความสามารถ และราคาที่ค่อนข้างประหยัดเมื่อเปรียบเทียบกับอุปกรณ์ตรวจจับภาพเชิงลึกอื่นๆ ทำให้ในปัจจุบันมีหลากหลายงานวิจัยให้ความสนใจเลือกใช้กล้อง Kinect ในการตรวจจับร่างกายมนุษย์ และแยกแยะการเคลื่อนไหวของมนุษย์ [10, 11]

3. วิธีการดำเนินการวิจัย

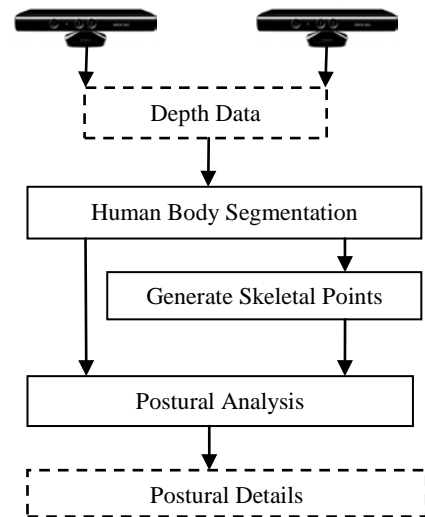
ในงานวิจัยนี้มีท่าทางที่ใช้ในการวิเคราะห์เพื่อแจกแจงท่าทางของผู้สูงอายุในระหว่างการรับชมโทรทัศน์ประกอบด้วย ท่ายืน ท่านั่ง และท่านอน โดยในท่านั่งนั้นจะแบ่งเป็นท่านั่งที่พื้น ท่านั่งบนเก้าอี้ และโซฟา และสามารถดูรายละเอียดเพิ่มเติมว่าในขณะที่ผู้สูงอายุกำลังนั่งนั้นเป็นการนั่งแบบงอหลังหรือ

เอนหลังอยู่ ในส่วนของท่าขึ้น และทำนั่งสามารถดูได้ว่า ผู้สูงอายุกำลังยกมือหรือเอามือพาดหัวอยู่หรือไม่ดังแสดงใน ตารางที่ 1 โดยในงานวิจัยนี้จะใช้กล้อง Kinect สองตัวในการ ตรวจจับท่าทางของผู้สูงอายุที่มีความถูกต้องมากยิ่งขึ้น ซึ่งแต่ละกล้องจะมีการทำงานในโหมดที่แตกต่างกัน โดยประกอบด้วย Seat Mode และ Default Mode โดย Default Mode มักพบว่าหลายๆครั้งมีการวิเคราะห์จุดต่างๆ บนร่างกาย ผิดพลาดในขณะที่นั่งพิงพนักพิงของโซฟาหรือเก้าอี้ โดยในงานวิจัยนี้จะนำ Seat Mode มาช่วยในการวิเคราะห์ ซึ่งพบว่า Seat Mode นั้นสามารถวิเคราะห์ท่าที่นั่งได้ดีกว่า Default Mode แต่ด้วยข้อมูลจุดบนร่างกายที่มีความละเอียดน้อยกว่าทำให้ไม่สามารถวิเคราะห์หลายๆท่าได้ จึงมีความจำเป็นที่จะต้อง วิเคราะห์แบบควบคู่กันไป โดยกล้อง Kinect หนึ่งตัวนั้น สามารถทำงานได้โหมดเดียวเท่านั้นจึงเป็นสาเหตุให้งานวิจัยนี้ เลือกใช้กล้องสองตัวในการพัฒนาระบบ

ตารางที่ 1: รายละเอียดท่าทางในงานวิจัย

ยืน	
นั่งเก้าอี้ และ โซฟา	
	
นั่งที่พื้น	
	
นอน	

ในงานวิจัยนี้จะใช้ทั้งข้อมูลภาพเชิงลึก และข้อมูลจุดต่าง ๆ บนร่างกายที่สามารถวิเคราะห์ออกมาได้จากโครงสร้างร่างกาย โดยขั้นตอนในการวิเคราะห์ท่าทางนั้นแสดงไว้ในภาพที่ 1 โดยจากรูปกรอบสี่เหลี่ยมเส้นประแทนความหมายชุดข้อมูล และกรอบสี่เหลี่ยมเส้นทึบแทนความหมายของฟังก์ชันที่ใช้ในการวิเคราะห์ข้อมูล



ภาพที่ 1: ขั้นตอนการจำแนกท่าทาง

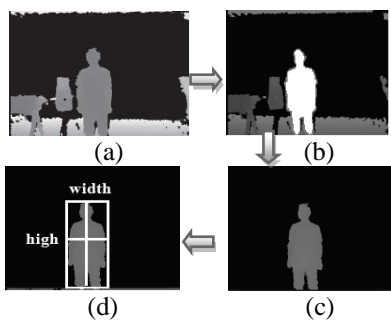
โดยภาพที่ 1 มีรายละเอียดในแต่ละขั้นตอนของการวิเคราะห์ ท่าทางดังต่อไปนี้

3.1 Human Body Segmentation

ในขั้นตอนนี้ทำการรับข้อมูลมาจากภาพเชิงลึกเพื่อนำมา แยกแยะส่วนที่มีความเป็นไปได้ที่จะเป็นร่างกายมนุษย์ซึ่งมี ขั้นตอนดังต่อไปนี้

1. ทำการดึงภาพเชิงลึกที่ถูกบันทึกโดยกล้อง Kinect ดัง แสดงในภาพ 2 (a)
2. ค้นหาส่วนที่มีความเป็นไปได้ที่จะเป็นร่างกายมนุษย์ เมื่อตรวจพบส่วนที่มีความเป็นไปได้ว่าเป็นร่างกาย มนุษย์แล้ว ให้ทำสัญลักษณ์โดยสีที่แตกต่างจากจุด อื่นๆไว้ดังแสดงในภาพ 2 (b)
3. ลบส่วนที่เป็นพื้นหลังทั้งหมดออกจากภาพเชิงลึกโดย แทนที่ด้วยจุดสีดำดังแสดงในภาพที่ 2 (c) หา อัตราส่วนความกว้าง และความสูงของส่วนที่ค้นพบว่า

เป็นร่างกายดังแสดงในภาพที่ 2 (d) เพื่อใช้ในการวิเคราะห์ท่าทางตามสมการที่ 1



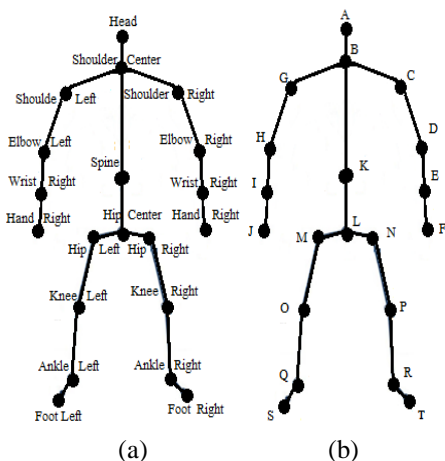
ภาพที่ 2: การแยกแยะร่างกายมนุษย์ออกจากภาพเชิงลึก

$$\text{Depth Data Ratio} = \frac{\text{Width}}{\text{Height}} \quad (1)$$

Width และ Height คือจุดที่สูง และกว้างที่สุดของภาพเชิงลึกที่ถูกระบุว่าเป็นร่างกายมนุษย์

3.2 Generate Skeletal Points

ในขั้นตอนนี้จะเป็นการนำภาพเชิงลึกมาหาจุดบนร่างกาย โดยงานวิจัยนี้เลือกใช้ Windows SDK Library Functions ซึ่งสามารถค้นหาจุดบนร่างกายได้ 2 แบบได้แก่ Seat Mode และ Default Mode โดย Default Mode สามารถค้นหาจุดบนร่างกายได้ 20 จุดทั่วร่างกาย และ Seat Mode สามารถค้นหาจุดได้ 10 จุดจากส่วนแขน และเหนือไหล่ดังแสดงในภาพที่ 3 (a) โดยในงานวิจัยจะแทนแต่ละจุดด้วยตัวอักษรดังแสดงในภาพที่ 3 (b)



ภาพที่ 3: องค์ประกอบของจุดต่างๆ บนร่างกาย

3.3 Postural Analysis

หลังจากที่เก็บข้อมูลซึ่งเป็นองค์ความรู้ในการวิเคราะห์ท่าทางได้ทั้งหมดแล้ว ข้อมูลทั้งหมดจะถูกนำมาวิเคราะห์ในขั้นตอนนี้ ซึ่งแต่ละท่าทางมีวิธีการในการวิเคราะห์ดังต่อไปนี้

3.3.1 การวิเคราะห์ท่านอน

ในท่านอนจะใช้อัตราส่วนที่เปรียบเทียบด้านกว้าง และด้านสูงจากภาพเชิงลึกในการวิเคราะห์ท่าทางการนอน ซึ่งค่าในการวิเคราะห์นั้นเป็นค่าเฉลี่ยที่บันทึกได้จาก 6 คนประกอบด้วยผู้ชาย 3 คนและผู้หญิง 3 คน ที่มีสัดส่วนและความสูงแตกต่างกันเพื่อให้มีความหลากหลายของข้อมูล โดยจากการวิเคราะห์ท่านอนแล้วพบว่าค่าอัตราส่วนของท่านอนมากกว่า 1.732

3.3.2 การวิเคราะห์ท่านั่งที่พื้น

ในงานวิจัยนี้จะนำข้อมูลมาวิเคราะห์ท่าที่นั่งที่พื้นต่อ โดยจะวิเคราะห์ท่าที่นั่งที่พื้นจากค่าเฉลี่ยของระยะห่างระหว่างสะโพกทั้งสองข้างไปยังพื้นดังแสดงในภาพที่ 4 การคำนวณระยะห่างระหว่างสะโพกด้านซ้ายไปยังพื้นถูกนำเสนอในสมการที่ 2 และใช้วิธีการเดียวกันสำหรับการคำนวณระยะห่างระหว่างสะโพกด้านขวาไปยังพื้น หลังจากนั้นหาค่าเฉลี่ยของระยะห่างระหว่างสะโพกทั้งสองข้างไปยังพื้นคำนวณได้โดยใช้สมการที่ 3 ซึ่งค่าเฉลี่ยจากคน 6 คนพบว่า ค่าระยะห่างจะต้องมีค่าน้อยกว่า 3.5 จึงถือว่าเป็นการนั่งที่พื้น



ภาพที่ 4: ระยะห่างระหว่างสะโพกกับพื้นในท่าต่าง ๆ

$$D_{\text{Hip_Left}} = \frac{AX + BY + CZ + D}{\sqrt{A^2 + B^2 + C^2}} \quad (2)$$

$$D_{\text{Hip_Avg}} = \frac{D_{\text{Hip_Left}} + D_{\text{Hip_Right}}}{2} \quad (3)$$

โดยค่า A, B, C, D เป็นค่าสัมประสิทธิ์ของระนาบซึ่งในที่นี้คือพื้น และ (X, Y, Z) จะเป็นพิกัดในจุด M (Hip Left)

3.3.3 การวิเคราะห์ท่าขึ้นและทำนั่ง

ในงานวิจัยนี้จะวิเคราะห์ท่าทำนั่ง และท่าขึ้นจากการงอของเข่า ซึ่งเมื่อนักวิจัยต้องการที่จะนั่งลงมของเข่าก็พับลงตามไปด้วย ในทางตรงกันข้ามในขณะที่นักวิจัยยืนอยู่มของหัวเข่าก็จะเปิดกว้างขึ้นตามไปด้วย ซึ่งในงานวิจัยนี้จะใช้สมการที่ 4 คำนวณการงอเข่า ซึ่งในกรณีที่นั่งเหยียดขาของเข่าจะเปิดเช่นเดียวกับการยืนจึงต้องมีการวิเคราะห์เพิ่มด้วยสมการที่ 5 เพื่อให้สามารถตรวจสอบท่าทำนั่งได้ถูกต้องมากยิ่งขึ้น

$$\theta = \cos^{-1} \left(\frac{-a + b + c}{2*b*c} \right) \quad (4)$$

$$\alpha = \cos^{-1} \left(\frac{|y_1 - y_2|}{a} \right) \quad (5)$$

โดยค่า a ในสมการมีค่าเท่ากับ \vec{NR} และ \vec{MQ} ค่า b ในสมการมีค่าเท่ากับ \vec{NP} และ \vec{MO} และ ค่า c ในสมการมีค่าเท่ากับ \vec{PR} และ \vec{OQ} และในส่วน y_1 และ y_2 คือค่าพิกัดในแนวแกน Y ของจุด N และ M

3.3.4 การวิเคราะห์การเอามือพาดหัว

ในการวิเคราะห์การเอามือพาดหัวนั้นในงานวิจัยนี้จะวิเคราะห์สองประเด็นด้วยกัน โดยประเด็นแรกคือเปรียบเทียบระหว่างข้อศอก และไหล่ โดยในขณะที่เอามือพาดหัวข้อศอกจะต้องอยู่เหนือไหล่ทั้งสองข้าง หากข้อศอกทั้งสองข้างอยู่เหนือไหล่จึงมีความเป็นไปได้ที่จะเป็นการเอามือพาดหัว ประเด็นที่สองคือการตรวจสอบมุมการงอของข้อศอกโดยจะคำนวณตามสมการที่ 4 โดย a ในสมการมีค่าเท่ากับ \vec{CE} และ \vec{GI} ค่า b ในสมการมีค่าเท่ากับ \vec{CD} และ \vec{GH} และ ค่า c ในสมการมีค่าเท่ากับ \vec{DE} และ \vec{HI}

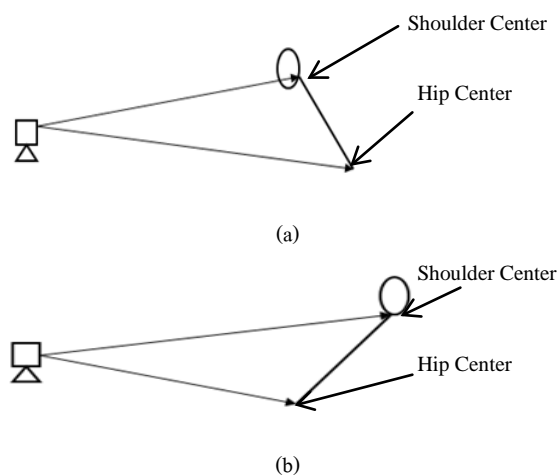
3.3.5 การวิเคราะห์ท่ายกมือ

สำหรับท่ายกมือในงานวิจัยนี้ให้ความสนใจเฉพาะในกรณีที่มืออยู่เหนือไหล่หรือทั้งมือ และข้อศอกอยู่เหนือไหล่ขึ้นไปเท่านั้น โดยในกรณีที่มือข้างใดข้างหนึ่งอยู่เหนือไหล่จะถือได้

ว่าข้างนั้นกำลังยกมือ แต่ในกรณีที่ทั้งมือ และข้อศอกอยู่เหนือไหล่แล้วนั้นจะต้องมีการตรวจสอบมุมการงอของข้อศอกเพิ่มเติมว่ามุมที่งอเป็นการเอามือพาดหัวหรือไม่ หากไม่ใช่จึงจะถือว่าทำนั้นเป็นการยกมือ

3.3.6 การวิเคราะห์ท่านั่งงอตัวและเอนตัว

ในการวิเคราะห์ท่างอตัว และเอนตัวนั้นในงานวิจัยให้ความสนใจในท่า นั่งเท่านั้น โดยจะวิเคราะห์โดยการเปรียบเทียบจากจุดสองจุดได้แก่ จุด B (Shoulder Center) และ จุด L (Hip Center) โดยดูจากพิกัดค่า Z ซึ่งบ่งบอกถึงระยะห่างระหว่างก้นไปยังจุดนั้น ๆ โดยหากพิกัด Z ของจุด B มีค่าน้อยกว่าจุด L จะมีความหมายว่าเป็นการงอตัวดังแสดงในภาพที่ 5 (a) และหากพิกัด Z ของจุด B มีค่ามากกว่าจุด L จะมีความหมายว่าเป็นการเอนตัวดังแสดงในภาพที่ 5 (b)



ภาพที่ 5: การนั่งงอตัว (a) การนั่งเอนตัว (b)

4. ผลการดำเนินงาน

ในงานวิจัยนี้ทดสอบความถูกต้องของการจำแนกแต่ละท่าทางโดยจำลองเหตุการณ์ขณะรับชมโทรทัศน์ โดยให้ผู้ถูกตรวจจับแสดงอาการต่าง ๆ ตามที่กำหนดไว้โดยใช้จำนวนผู้ทดสอบ 8 คน ซึ่งเป็นคนละกลุ่มกับกลุ่มคนที่ใช้ในการบันทึกข้อมูลท่าทางซึ่งมีทั้งบุคคลที่มีลักษณะรูปร่างคล้ายและแตกต่างจากกลุ่มตัวอย่างที่ใช้ในการบันทึกข้อมูล เพื่อทดสอบถึงความยืดหยุ่นของการแยกแยะท่าทางในงานวิจัย โดยผลลัพธ์ของแต่ละท่าทางได้แสดงไว้ในตารางที่ 2

ตารางที่ 2: ความถูกต้องในการตรวจจับแต่ละท่าทาง

ท่าหลัก	ท่าย่อย	ความถูกต้อง(%)
ยืน	-	93.50
	ยกมือข้างซ้าย	87.85
	ยกมือข้างขวา	87.55
	ยกมือสองข้าง	88.61
นั่งบนเก้าอี้และโซฟา	-	91.81
	งอตัว	89.81
	เอนตัว	90.20
	ยกมือข้างซ้าย	87.60
	ยกมือข้างขวา	88.36
	ยกมือสองข้าง	88.89
	เอามือพาดหัว	88.29
	-	90.78
นั่งบนพื้น	งอตัว	89.03
	เอนตัว	89.87
	ยกมือข้างซ้าย	88.70
	ยกมือข้างขวา	87.98
	ยกมือสองข้าง	89.13
	เอามือพาดหัว	88.45
	-	90.23
	นอนบนพื้นและโซฟา	-

5. สรุป

จากการทดสอบการตรวจจับท่าทางพบว่ากล้อง Kinect สามารถจำแนกท่าทางได้ดี แม้ว่าผู้ทดสอบจะมีรูปร่างและความสูงแตกต่างกันออกไป โดยความผิดพลาดในการตีความหมายของท่าทางนั้นเกิดจาก 2 ประเด็นด้วยกัน ได้แก่ การบ่งบอกจุดบนร่างกายที่ผิดพลาดโดย Library Function และการตีความที่ผิดพลาดในระหว่างการเปลี่ยนแปลงท่าทางด้วยสาเหตุจากค่าที่ใช้ในการแบ่งแยกท่าทางอาจไม่มีความเหมาะสมกับบางบุคคล ดังนั้นหากงานวิจัยนี้ไปประยุกต์ใช้กับหลักการเหมืองข้อมูล (Data Mining) ร่วมกับระบบปัญญาประดิษฐ์ (Artificial Intelligence) ก็จะช่วยให้การตีความหมายของระบบเพิ่มความถูกต้องมากยิ่งขึ้น และหากนำระบบนี้ไปพัฒนาร่วมกับระบบรู้จำการแสดงออกทางใบหน้าก็จะสามารถช่วยให้การตีความพฤติกรรมมีความถูกต้องและแม่นยำมากขึ้นอีกด้วย

เอกสารอ้างอิง

- [1] United Nations, "World Population Ageing 1950-2050," [Online] Available: www.un.org/esa/population/publications/worldageing19502050/, (Accessed 12 November 2013).
- [2] S. Kensaku, "Older People and Television Viewing in Japan," *Journal of NHK Broadcasting Studies* 2010, No. 8, pp. 63-94, 2010.
- [3] K. Reid, "Lifeline or Leisure? : TV's Role in the Lives of the Elderly," [Online] Available: www.medialit.org/reading-room/lifeline-or-leisure-tvvs-role-lives-elderly/, (Accessed 8 December 2013).
- [4] Mann W. C., "The Aging Population and Its Needs," *IEEE Journals & Magazines, Pervasive Computing*, Vol.3, Issue 2, pp. 12-14, 2004.
- [5] M. Yu, S. M. Naqvi and J. Chambers, "Fall detection in the elderly by head tracking," *Statistical Signal Processing, 2009. SSP '09. IEEE/SP 15th Workshop on*, Wales, United Kingdom, Aug 31 - Sept 3, 2009.
- [6] J. T. Kim, J. Y. Soh, S. H. Kim, and K. Y. Chung, "Emergency Situation Alarm System Motion using Tracking of People Like Elderly Live Alone," *Information Science and Applications (ICISA), 2013 International Conference on*, Pattaya, Thailand, Jun 24-26, 2013.
- [7] F. Cardile, G. Iannizzotto and F. La Rosa, "A Vision-Based System for Elderly Patients Monitoring," *Human System Interactions (HSI), 2010 3rd Conference on*, Rzeszow, Poland, May 13-15, 2010.
- [8] M. Fahim, I. Fatima, S. Lee and Y. K. Lee, "Daily Life Activity Tracking Application for Smart Homes using Android Smartphone," *Advanced Communication Technology (ICACT), 2012 14th International Conference on*, PyeongChang, South Korea, Feb 19-22, 2012.
- [9] A. Correa, A. Morell, M. Barcelo and J. L. Vicario, "Navigation system for elderly care applications based on wireless sensor networks," *Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European*, Bucharest, Romania, Aug 27-31, 2012.
- [10] O. Patsadu, C. Nukoolkit, and B. Watanapa, "Human gesture recognition using Kinect camera," *The 9th International Joint Conference on Computer Science and Software Engineering (JCSSE 2012)*, Bangkok, Thailand, May 30- Jun 1, 2012.
- [11] Y. Liu, Z. Zhang, A. Li and M. Wang, "View independent human posture identification using Kinect," *Biomedical Engineering and Informatics (BMEI), 2012 5th International Conference on*, Chongqing, China, Oct 16-18, 2012.

การปรับปรุงการใช้เซลล์ลาร์ออโตมาตาแบบพื้นฐานในการเข้ารหัสรูปภาพ

An Improved Method for Elementary Cellular Automata Image Encryption Scheme

มงคล ทองไกรแก้ว (Mongkol Thongkraikaew)¹ และรุ่งรัตน์ เวียงศรีพานาวลัย (Rungrat Wiangsripanawan)²

^{1,2}สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

mongkol_ttm@hotmail.com¹, kwrungra@kmitl.ac.th²

บทคัดย่อ

งานวิจัยนี้ปรับปรุงขั้นตอนการเข้ารหัสรูปภาพด้วยวิธีการของเซลล์ลาร์ออโตมาตาพื้นฐาน (Elementary Cellular Automata) เพื่อแก้ปัญหาให้กับภาพที่เข้ารหัสด้วยวิธีการเดิมให้ผลที่ไม่ดี เช่น ภาพที่เข้ารหัสแล้วไม่สามารถปกปิดเค้าโครงของภาพเดิมและภาพที่เข้ารหัสแล้วฮิสโตแกรมให้ค่าที่ไม่สม่ำเสมอ โดยเพิ่มขั้นตอนการพรีโพรเซส (Pre-process) เพื่อให้ค่าของแต่ละพิกเซลก่อนนำไปเข้ารหัสให้มีความกระจายมากขึ้น อีกทั้งมีการนำหลักการการเปลี่ยนภาพขาวดำเป็นภาพสีเทา มาใช้ร่วมกับการเข้ารหัสที่ปรับปรุงขึ้นมาใหม่ ทำให้สามารถเข้ารหัสภาพขาวดำซึ่งแต่เดิมไม่สามารถทำได้ ผลการทดลองแสดงให้เห็นว่าการเข้ารหัสแบบใหม่สามารถเข้ารหัสภาพได้ทุกประเภททุกรูปแบบและปกปิดข้อมูลได้ดีกว่าเดิม

คำสำคัญ: เข้ารหัสรูปภาพ เซลล์ลาร์ออโตมาตา ภาพขาวดำ

Abstract

This paper presents a method to improve the existing Elementary Cellular Automata Image Encryption Scheme that has some problems. First, some cipher images keep their traces such as the border of the heads or bodies which can be seen with bare eyes. Second, the histograms of some cipher images are not distributed well. Third, it cannot encrypt the black-and-white images. We improve the randomness of the cipher image by adding a pre-processing step to each pixel before encryption. We also apply the method to converse a black-and-white image to a grey-scale image and then use our improved scheme to encrypt it. The experimental results show that with our improvement, all types of images (color, grey-scale and

black-and-white) can be encrypted and a better concealment is obviously provided.

Keyword: Image Encryption, Elementary Cellular Automata, Black-and-White Image.

1. บทนำ

ปัจจุบันมีการใช้งานรูปภาพอย่างกว้างขวางทั้งในโซเชี่ยลเน็ตเวิร์คและสื่อมัลติมีเดีย ข้อมูลรูปภาพส่วนใหญ่ถูกแบ่งปันผ่านเครือข่ายอินเทอร์เน็ตซึ่งไม่ปลอดภัย จึงมีการนำวิธีการของการเข้ารหัสแบบสมมาตร (Symmetric Key Cryptography) เช่น อัลกอริทึม AES มาใช้เพื่อเพิ่มความปลอดภัย แต่การที่ AES เข้ารหัสแบบบล็อก (Block Cipher) ซึ่งไม่เหมาะกับการเข้ารหัสรูปภาพที่มีขนาดใหญ่และต้องการความเร็ว การเข้ารหัสรูปภาพเหมาะกับวิธีการเข้ารหัสแบบสตรีม (Stream Cipher) ซึ่งเป็นการเข้ารหัสแบบสมมาตรอีกประเภทหนึ่ง ที่เน้นความเร็วอย่างไรก็ดี การนำอัลกอริทึมสำหรับเข้ารหัสข้อมูลทั่วไปทางสตรีม เช่น RC4 ซึ่งไม่ได้ถูกออกแบบมาให้เข้ารหัสข้อมูลรูปภาพมาเข้ารหัสไฟล์รูปภาพอาจจะให้ผลลัพธ์ที่ไม่ดี เพราะข้อมูลหรือพิกเซลของรูปภาพมีความซ้ำซ้อนสูงกว่าข้อมูลประเภทตัวอักษรมาก เมื่อเข้ารหัสแล้วอาจทำให้มีเค้าโครงของภาพต้นฉบับหลงเหลืออยู่ หรือให้ค่าฮิสโตแกรมที่ไม่กระจาย

ปัจจุบันมีผู้เสนอวิธีเข้ารหัสรูปภาพหลากหลายวิธี เซลล์ลาร์ออโตมาตาเป็นหนึ่งในนั้นมีการนำเซลล์ลาร์ออโตมาตาทั้งแบบ 2 มิติ และ 1 มิติมาประยุกต์ใช้ในการเข้ารหัส สำหรับในงานวิจัยนี้ได้นำงานวิจัยของ [1,2] มาปรับปรุง ซึ่งใช้เซลล์ลาร์ออโตมาตาแบบ 1 มิติ ในการสร้างคีย์และการเข้ารหัส ข้อดีของวิธีการนี้คือมีความซับซ้อนน้อยกว่าวิธีการอื่น และใช้เวลาน้อย อย่างไรก็ตามก็ดีจากการทดลองกับภาพในฐานข้อมูล USC-SIPI

[3] พบว่าการเข้ารหัสแบบเดิมใน [1] นั้นให้ผลที่ไม่ค่อยดีมากกับภาพที่มีพิกเซลโทนสีเดียวกันในจำนวนมากๆ เช่น ภาพห้องฟ้า ภาพทะเล เนื่องจากภาพที่ได้จากการเข้ารหัสจะเห็นเค้าโครงของภาพเดิม และการเข้ารหัสด้วยวิธีการเดิมนั้นไม่สามารถเข้ารหัสภาพขาวดำ (Black and White) ได้

ดังนั้นในงานวิจัยนี้จึงทำการเพิ่มขึ้นตอนก่อนที่จะนำข้อมูลของแต่ละพิกเซลไปเข้ารหัส โดยมีวัตถุประสงค์ในการลดความซ้ำซ้อนของค่าของพิกเซลตั้งต้น และ มีการนำหลักการของ [6] มาประยุกต์ ในการเปลี่ยนภาพขาวดำ ให้เป็นภาพสีเทา แล้วจึงนำภาพสีเทานำไปเข้ารหัส

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 เซลลูลาร์ออโตมาตา

2.1.1 เซลลูลาร์ออโตมาตาพื้นฐาน

เซลลูลาร์ออโตมาตา [2,4] เป็นแบบจำลองทางคณิตศาสตร์ที่ใช้ในการอธิบายระบบต่างๆ ที่สามารถแยกองค์ประกอบเป็นส่วนๆ และเกี่ยวข้องกับเวลาที่เปลี่ยนแปลงแบบไม่ต่อเนื่อง (discrete) เรียกแต่ละส่วนว่าเซลล์ (cell) แต่ละเซลล์มีสถานะ (state) ของตัวเองและเปลี่ยนแปลงเมื่อเวลาเปลี่ยน t ไป $t+1$ สถานะใหม่ถูกกำหนดโดยสถานะของย่านข้างเคียง (neighborhood) ของเซลล์ ณ เวลา t โดยในแต่ละเซลล์มีฟังก์ชันหรือเรียกว่ากฎ (rule) เป็นตัวกำหนดสถานะใหม่ โดยการเปลี่ยนสถานะสามารถคำนวณได้ดังสมการที่ (1)

$$s_i^{t+1} = f_i(s_{i\text{neighborhood}}^t) \tag{1}$$

ให้ s_i^{t+1} แทนค่าสถานะของเซลล์ i ที่เวลา $t+1$, f_i แทนฟังก์ชันหรือกฎของตำแหน่ง i $s_{i\text{neighborhood}}^t$ แทนค่าสถานะของย่านข้างเคียงที่เวลา t

เซลลูลาร์ออโตมาตาประเภท 1 มิติ มีค่าสถานะของแต่ละเซลล์เป็นได้เพียง 0 และ 1 เท่านั้น กำหนดให้สถานะย่านข้างเคียงของเซลล์คือ $s_{i-1}^t, s_i^t, s_{i+1}^t$ สถานะใหม่ของเซลล์ i จะมีค่าดังสมการที่ (2)

$$s_i^{t+1} = f_i(s_{i-1}^t, s_i^t, s_{i+1}^t) \tag{2}$$

จากสมการที่ (2) เห็นว่าสถานะย่านข้างเคียงมี 3 ค่า แสดงว่ารูปแบบสถานะย่านข้างเคียงมีทั้งหมด $2^3 = 8$ รูปแบบ และมีกฎ

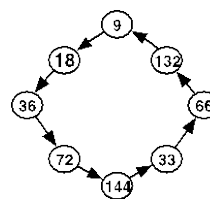
ได้ทั้งหมด $2^8 = 256$ กฎ ยกตัวอย่างการเปลี่ยนสถานะใหม่ในแต่ละกฎในตารางที่ 1

ตารางที่ 1: กฎของเซลลูลาร์ออโตมาตาแบบพื้นฐาน

ตำแหน่ง	7	6	5	4	3	2	1	0
สถานะย่านข้างเคียง	111	110	101	100	011	010	001	000
กฎ 90 (01011010)	0	1	0	1	1	0	1	0
กฎ 150 (10010110)	1	0	0	1	0	1	1	0
กฎ 255 (11111111)	1	1	1	1	1	1	1	1

2.1.2 แอทแทรกเตอร์ (Attractor)

เซลลูลาร์ออโตมาตาพื้นฐาน มีจำนวนเซลล์จำนวน 8 เซลล์เท่านั้น ดังนั้นการหาค่าสถานะใหม่ของเซลล์ที่ 1 และ 8 จะไม่มีสถานะย่านข้างเคียง s_{i-1}^t และ s_{i+1}^t ตามลำดับ จึงได้มีการนำเงื่อนไขแบบคาบ (periodic boundary) มาประยุกต์ใช้ โดยตั้งค่าสถานะใหม่ของเซลล์ที่ 1 ให้มีสถานะย่านข้างเคียงคือสถานะของเซลล์ที่ 8 และตั้งค่าสถานะย่านข้างเคียงของเซลล์ที่ 8 คือค่าสถานะของเซลล์ที่ 1



ภาพที่ 1: แผนภาพการเปลี่ยนสถานะแอทแทรกเตอร์ของกฎที่ 2

ดังนั้นเมื่อมีการเปลี่ยนสถานะไปเรื่อยๆ ค่าของทุกเซลล์จะวนกลับมาเท่ากับค่าสถานะเริ่มต้น เรียกคุณสมบัติแบบนี้ว่าแอทแทรกเตอร์ (Attractor) [2] สามารถเขียนเป็นแผนภาพตัวอย่างได้ตามภาพที่ 1 ในการคำนวณหาแอทแทรกเตอร์คำนวณจากกฎและสถานะเริ่มต้น ซึ่งจะมีบางสถานะในบางกฎเท่านั้นที่จะมีคุณสมบัติเป็นแอทแทรกเตอร์ได้ ซึ่งเมื่อนำสถานะทั้งหมดของแอทแทรกเตอร์ มา exclusive-or (xor) ค่าที่ได้จะเท่ากับ 0 สามารถเขียนเป็นสมการได้ดังนี้

$$st(1) \oplus st(2) \oplus \dots \oplus st(k) = 0 \tag{3}$$

ให้ $st(1), st(2)$ คือค่าสถานะของช่วงเวลาที่ 1 และ 2, k แทนจำนวนสถานะทั้งหมดของแอทแทรกเตอร์ และ \oplus แทนตัวดำเนินการ xor

2.2 การเข้ารหัสรูปภาพ

ในหลายบทความวิจัยมีการนำทฤษฎีต่างๆ มาประยุกต์ เพื่อเสนอวิธีการเข้ารหัสที่เหมาะสมสำหรับเข้ารหัสรูปภาพ โดยเฉพาะ [5-6] วิธีที่ได้รับความนิยมเป็นอย่างมากคือวิธีที่ Zhang และคณะ[7] นำหลักการ Chaotic system โดยใช้ Chaotic map ช่วยในการสลับข้อมูลที่ซ้ำกันให้เกิดความวุ่นวายของข้อมูลจนทำให้ไม่เหลือเค้าโครงของวัตถุหลงเหลืออยู่ อีกวิธีหนึ่งคือการสร้างสตรีมไซเฟอร์สำหรับเข้ารหัสรูปภาพ โดยเฉพาะ ซึ่งแต่ละงานวิจัยมีการนำเสนอวิธีการเข้ารหัสและสร้างคีย์ที่แตกต่างกันเช่น ใช้ Ant Colony Optimization (ACO) ในการสร้างคีย์สำหรับเข้ารหัส [8] เป็นต้น

2.3 การเข้ารหัสรูปภาพโดยใช้เซลล์ลาร์อโตมาตา

Wolfram [9] ศึกษาและริเริ่มการเข้ารหัสข้อมูลรูปภาพโดยใช้หลักการของเซลล์ลาร์อโตมาตา Chen และ Lai [10-11] นำเซลล์ลาร์อโตมาตาแบบ 2 มิติมาใช้ในการสร้างตัวเลขสุ่มเทียม (Pseudorandom Number Generator) Li และคณะ [12] นำเซลล์ลาร์อโตมาตาแบบ 2 มิติ มาใช้ในขั้นตอนการเข้ารหัส

สำหรับเซลล์ลาร์อโตมาตาแบบพื้นฐานนั้น Jun[2] นำคุณสมบัติพิเศษในการเปลี่ยนสถานะของแอทแทรกเตอร์ในสมการที่ (3) มาประยุกต์ใช้ดังนี้

$$plain \oplus st(1) \oplus st(2) \oplus \dots \oplus st(t) = cipher \quad (4)$$

$$cipher \oplus st(t+1) \oplus \dots \oplus st(k) = plain \quad (5)$$

ให้ $plain$ คือค่าของพิกเซลก่อนทำการเข้ารหัส $cipher$ คือค่าของพิกเซลหลังทำการเข้ารหัสแล้ว t แทนจำนวนสถานะที่น้อยกว่าสถานะทั้งหมดและ k แทนจำนวนสถานะทั้งหมด

วนิดา และ นันทิกา [1] นำงานวิจัยของ Jun มาปรับปรุงโดยนำวิธีการของตัวเลขสุ่มเทียมมาเพิ่มขนาดของจำนวนคีย์สเปซให้เพิ่มขึ้นจำนวนมากขึ้น และ ปรับปรุงวิธีการในการเข้ารหัสให้สามารถเข้ารหัสได้ทั้งภาพสีและภาพสีเทา (Gray scale)

2.4 ปัญหาของการเข้ารหัสด้วยวิธีการของวนิดา

วิธีการเข้ารหัสของวนิดามีข้อเสียคือ ภาพบางภาพเช่น ภาพท้องฟ้าและภาพที่ 4(a) เมื่อเข้ารหัสแล้วคงเหลือเค้าโครงบางอย่างของภาพต้นฉบับ และภาพบางภาพเมื่อเข้ารหัสแล้ว

ฮิสโตแกรมของภาพที่ผ่านการเข้ารหัสให้ค่าไม่สม่ำเสมอ และวิธีการนี้ไม่สามารถเข้ารหัสภาพขาวดำได้

3. วิธีการเข้ารหัสภาพแบบใหม่

เนื่องจากการวิจัยนี้นำวิธีการของวนิดามาปรับปรุง จึงขอกล่าวถึงวิธีการของวนิดาดังนี้

3.1 วิธีการเข้ารหัสของเดิม

งานวิจัยของวนิดา คีย์ที่ใช้ในการเข้ารหัสคือ (rule, seedstate, seedtime) rule แทนกฎที่ใช้ในการเข้ารหัส seedstate แทนค่าเริ่มต้นที่ใช้ในการหาตัวเลขสุ่มเทียมเพื่อหาสถานะเริ่มต้นและ seedtime แทนจำนวนสถานะที่ใช้ในการเข้ารหัส ตารางที่ 2 แสดงการเข้ารหัสแบบเดิมดังนี้

สำหรับการถอดรหัสต้นั้น จะมีวิธีเหมือนกับการเข้ารหัส โดยเริ่มจากคำนวณหาค่าเริ่มต้นต่างๆจากคีย์ จากนั้นนำพิกเซลที่จะถอดรหัสมาสลับบิตเช่นเดียวกับการเข้ารหัส เมื่อได้ค่าทั้งหมดให้ทำการถอดรหัสด้วยสมการที่ 8 ดังนี้

$$p_pic(r,c) = c_pix(r,c) \oplus st(t+1,r,c) \oplus \dots \oplus st(k,r,c) \quad (8)$$

ตารางที่ 2: วิธีการเข้ารหัสแบบเดิม

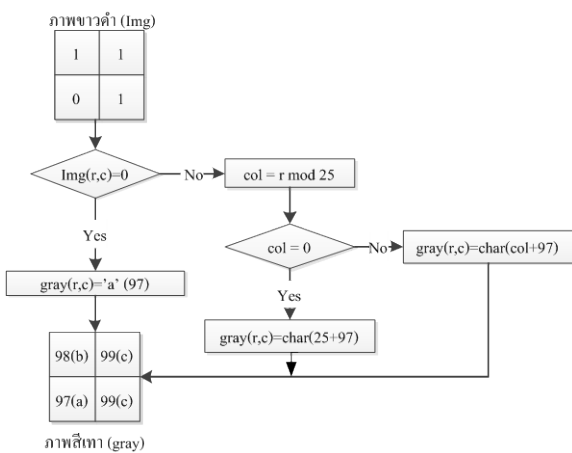
- | |
|--|
| <ol style="list-style-type: none"> 1) นำกฎมาหาค่าของแอทแทรกเตอร์เก็บไว้ในอาร์เรย์ P 2) สุ่มสถานะเริ่มต้นจาก seedstate เก็บไว้ในอาร์เรย์ S 3) สุ่มจำนวนสถานะที่ใช้ในการเข้ารหัสจาก seedtime เก็บไว้ในอาร์เรย์ T 4) กำหนดสถานะเริ่มต้นให้กับแต่ละพิกเซลด้วยสมการ $st(1,r,c) = P((S(r,c) \bmod q) + 1) \quad (6)$ 5) เข้ารหัสแต่ละพิกเซล $p_pix(r,c) \oplus st(1,h,v) \oplus \dots \oplus st(t,h,v) = c_pix \quad (7)$ $t = T(r,c) \bmod k - 1$ 6) สลับบิต |
|--|

ให้ $st(i,r,c)$ แทนสถานะที่ i ของพิกเซลตำแหน่งแถว r และหลัก c , $S(r,c)$ แทนค่าตัวเลขสุ่มที่อยู่ในอาร์เรย์ S ที่แถว r และหลัก c , $P(i)$ แทนค่าสถานะของอาร์เรย์ P ที่ตำแหน่ง i , q แทนจำนวนสถานะทั้งหมดที่อยู่ในอาร์เรย์ P, $p_pix(r,c)$ แทนพิกเซลเพลนที่กซ์ที่ตำแหน่งแถว r และหลัก c , c_pix แทนพิกเซลไซเฟอร์ที่กซ์ที่ตำแหน่งแถว r และหลัก c และ $T(r,c)$ แทนค่าตัวเลขสุ่มที่อยู่ในอาร์เรย์ T

3.2 วิธีการเข้ารหัสแบบใหม่

วิธีการเข้ารหัสแบบใหม่ปรับปรุงข้อเสียของวิธีเข้ารหัสแบบเดิม โดยปรับปรุงอัลกอริทึมในการเข้ารหัสให้สามารถเข้ารหัสภาพได้ทุกประเภท และเพิ่มประสิทธิภาพการปกปิดข้อมูล โดยเพิ่มขั้นตอนสองขั้นตอนคือ หนึ่ง ขั้นตอนการเปลี่ยนจากภาพขาวดำเป็นภาพสีเทาในกรณีที่มีรูปภาพที่เข้ารหัสเป็นภาพขาวดำ ขั้นตอนที่สอง ในส่วนของอัลกอริทึมในการเข้ารหัส มีการทำพรีโปรเซสโดยการนำพิกเซลของ เฟลนเท็กซ์ค่านั้นไป xor กับพิกเซลไซเฟอร์ก่อนหน้า ก่อนที่จะนำพิกเซลนั้นไปเข้ารหัส

แผนผังในภาพที่ 3 แสดงขั้นตอนในการเข้ารหัสแบบใหม่ ซึ่งสังเกตได้ว่า ไม่ว่าจะป็นภาพสี ภาพสีเทา และ ภาพขาวดำ วิธีการในการเข้ารหัสเหมือนกันทั้งหมดแต่ภาพขาวดำจะมีการแปลงภาพขาวดำเป็นภาพสีเทาก่อน ซึ่งงานวิจัยนำวิธีการของ Sreelaja [6] มาประยุกต์ในการเปลี่ยนภาพขาวดำเป็นภาพสีเทา ซึ่งขั้นตอนในการเปลี่ยนภาพขาวดำเป็นภาพสีเทาสามารถดูได้ในภาพที่ 2



ภาพที่ 2: ขั้นตอนการเปลี่ยนภาพขาวดำเป็นภาพสีเทา [6]

หลังจากได้ภาพเรียบร้อยแล้ว (ภาพสี ภาพสีเทา ภาพสีที่ถูกแปลงมาจากภาพขาวดำ) ก่อนที่จะนำแต่ละพิกเซลไปเข้ารหัสในวิธีการเดิมให้นำพิกเซลนั้นไปทำพรีโปรเซสในสมการที่ (9) ก่อนนำไปเข้ารหัสในสมการที่ (10) และสิ้นสุดด้วยการสลับบิตตามแบบการเข้ารหัสแบบเดิม

$$pre_pix(r,c) = p_pix(r,c) \oplus c_pix(r,c - 1) \tag{9}$$

ถ้า $p_pix(r,0) : i = \{1...m\}$ ค่า $c_pix(r,c - 1) = 0$

$$c_pix(r,c) = pre_pix(r,c) \oplus st(1,r,c) \oplus ... \oplus st(t,r,c) \tag{10}$$

เมื่อ $pre_pix(r,c)$ แทนค่าสถานะก่อนทำการ xor กับสถานะในการเข้ารหัส

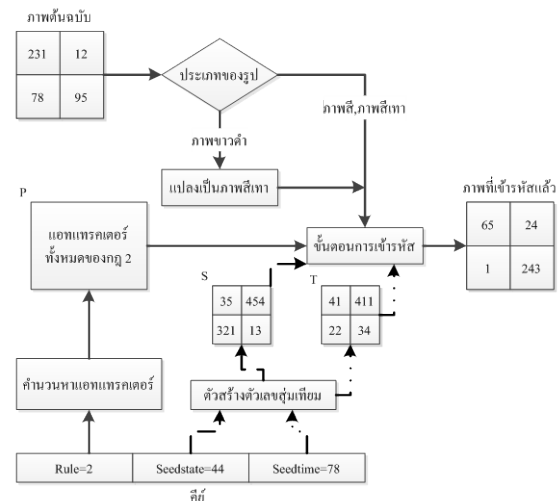
ขั้นตอนของการถอดรหัสมีวิธีการเช่นเดียวกับการถอดรหัสวิธีเดิม โดยนำพิกเซลไซเฟอร์เท็กซ์ไปสลับบิตและทำการถอดรหัสตามปกติดังสมการที่ (2) จากนั้นให้นำค่านั้นไป xor กับค่าของพิกเซลไซเฟอร์ก่อนหน้า จึงจะได้ค่าพิกเซล เฟลนเท็กซ์ดังสมการที่ (12) เมื่อทำงานครบทุกพิกเซลจะได้ภาพต้นฉบับคืนมา ยกเว้นในกรณีที่ภาพต้นฉบับเป็นภาพขาวดำ ให้ทำวิธีการย้อนกลับสมการในการแปลงภาพสีเทาเป็นภาพขาวดำ

$$pre_pix(r,c) = c_pix(r,c) \oplus st(c+1,r,c) \oplus ... \oplus st(k,r,c) \tag{11}$$

$$p_pix(r,c) = pre_pix(r,c) \oplus c_pix(r,c - 1) \tag{12}$$

ถ้า $c_pix(r,0) : i = \{1...m\}$ ค่า $c_pix(r,c - 1) = 0$

เมื่อ $pre_pix(r,c)$ แทนค่าสถานะก่อนทำการ xor กับสถานะในการถอดรหัส



ภาพที่ 3: ขั้นตอนการเข้ารหัสวิธีใหม่

4. ผลการทดลอง

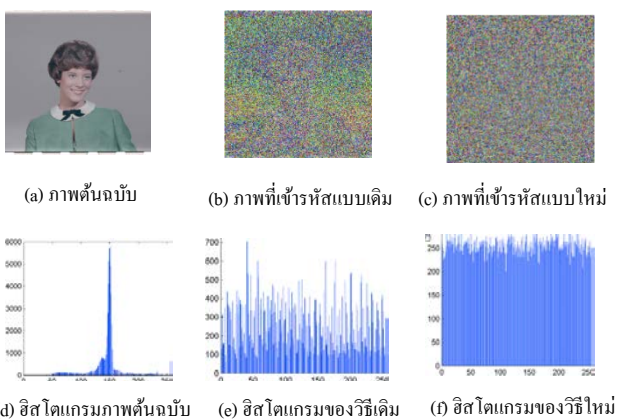
รูปภาพที่ใช้ในการทดสอบในงานวิจัยนี้ นำมาจากฐานข้อมูล USC-SIPI [3] ประกอบด้วยภาพสีและภาพสีเทา ในส่วนของภาพขาวดำนั้น ได้ทำการแปลงมาจากภาพสีเทาบางส่วน ขนาดของพิกเซลของภาพที่ใช้มี 3 ขนาดคือ 256×256 , 512×512 และ 1024×1024 ตามลำดับ ในการทดลองใช้

กฎเดียวกันเข้ารหัสในการเข้ารหัสแบบเก่าและแบบใหม่ ทดสอบโดยคุณสมบัติต่อไปนี้

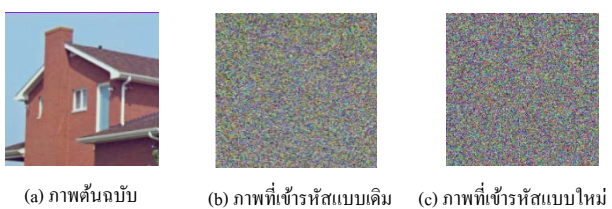
4.1 การปกปิดข้อมูล (Information Concealing)

วิธีการเข้ารหัสที่ดีไม่ควรเหลือเค้าโครงของภาพเดิมและค่าฮิสโตแกรมของภาพที่เข้ารหัสแล้วควรจะมีการกระจายอย่างสม่ำเสมอ ซึ่งการเข้ารหัสของงานวิจัยเดิมให้ผลไม่ดีกับภาพที่มีค่าพิกเซลบางพิกเซลจำนวนมาก เช่น ภาพที่ 4(b) เมื่อดูฮิสโตแกรมของภาพต้นฉบับ (ภาพ 4(d)) สังเกตได้ว่าค่าพิกเซลของสีน้ำเงินที่บริเวณตรงกลางมากกว่าพิกเซลอื่นมาก เมื่อนำค่านี้ไปเข้ารหัสจะได้ผลลัพธ์ที่มีค่าซ้ำกันในจำนวนมาก ทำให้เค้าโครงของภาพเดิมคงอยู่ (ภาพ 4(b)) ยิ่งไปกว่านั้นฮิสโตแกรมของภาพที่เข้ารหัสแล้ว ให้ค่าที่ไม่กระจายแต่เมื่อนำภาพ 4(a) ไปเข้ารหัสด้วยวิธีการใหม่พบว่าภาพที่เข้ารหัสแบบใหม่ไม่หลงเหลือเค้าโครงของภาพที่สามารถเห็นด้วยตาเปล่าและฮิสโตแกรมให้ผลดีกว่าภาพเดิมมากดังภาพ 4(f) ในภาพที่ 5 พบว่าภาพที่ได้จากการเข้ารหัสแบบเดิมและแบบใหม่ไม่สามารถจำแนกความแตกต่างได้ด้วยตาเปล่าแต่เมื่อดูด้วยฮิสโตแกรมในภาพ 5(e) และ 5(f) พบว่าการเข้ารหัสแบบใหม่ให้ผลดีกว่าการเข้ารหัสแบบเดิมค่อนข้างมาก

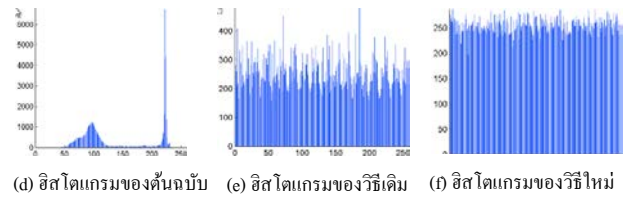
ฮิสโตแกรมในภาพที่ 6 และ 7 แสดงให้เห็นว่าแม้การเข้ารหัสแบบเดิมให้ผลของฮิสโตแกรมค่อนข้างดี แต่เมื่อนำมาเข้ารหัสด้วยวิธีการใหม่พบว่าการเข้ารหัสแบบใหม่ให้ผลที่ดีกว่าเสมอ



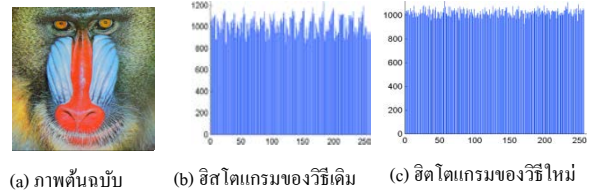
ภาพที่ 4: เปรียบเทียบภาพและฮิสโตแกรมของทั้งสองวิธี



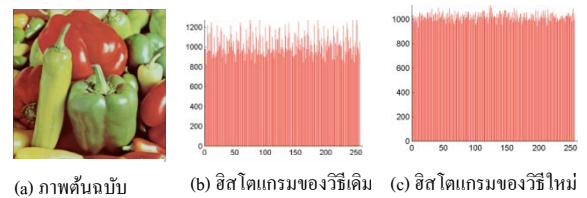
(a) ภาพต้นฉบับ (b) ภาพที่เข้ารหัสแบบเดิม (c) ภาพที่เข้ารหัสแบบใหม่



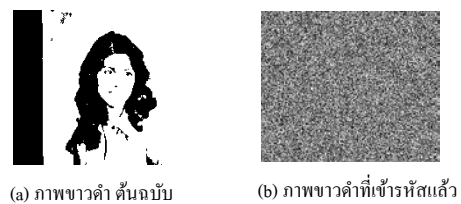
ภาพที่ 5: เปรียบเทียบภาพและฮิสโตแกรมของทั้งสองวิธี



ภาพที่ 6: เปรียบเทียบฮิสโตแกรมของทั้งสองวิธี



ภาพที่ 7: เปรียบเทียบฮิสโตแกรมของทั้งสองวิธี



ภาพที่ 8: ผลการเข้ารหัสภาพขาวดำ

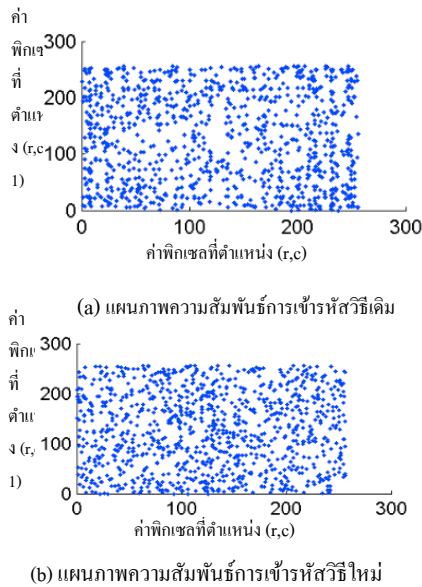
จากผลการทดลองสรุปได้ว่าการเข้ารหัสด้วยวิธีการใหม่ปกปิดข้อมูลได้ดีกว่าเดิมในภาพทุกประเภท

4.2 ความสัมพันธ์ระหว่างพิกเซล (Correlation Coefficient)

ในการทดสอบคุณสมบัติ Diffusion และ Confusion นั้น ทดสอบด้วยการคำนวณค่าความสัมพันธ์ระหว่าง 2 พิกเซลที่อยู่ติดกันของภาพที่เข้ารหัสแล้ว โดยสุ่มคู่พิกเซลที่อยู่ติดกันจำนวน 1000 คู่ ทั้งในแนวนอน แนวตั้ง และแนวเฉียง มาคำนวณหาค่าความสัมพันธ์ดังสมการต่อไปนี้

$$C_r = \frac{(N \sum_{j=1}^N x_j y_j - \sum_{j=1}^N x_j \sum_{j=1}^N y_j)}{(N \sum_{j=1}^N x_j^2 - (\sum_{j=1}^N x_j)^2) (N \sum_{j=1}^N y_j^2 - (\sum_{j=1}^N y_j)^2)} \quad (13)$$

เมื่อ x_j และ y_j แทนค่าพิกเซลที่อยู่ติดกันและ N แทนจำนวนคู่พิกเซลทั้งหมดที่เลือกมาใช้ในการคำนวณ



ภาพที่ 9: แผนภาพความสัมพันธ์ระหว่างสองพิกเซลทั้งสองวิธี

จากภาพที่ 9 สรุปได้ว่าการเข้ารหัสด้วยวิธีการใหม่มีความสัมพันธ์ของ 2 พิกเซลที่อยู่ติดกันน้อยกว่าการเข้ารหัสด้วยวิธีเดิม โดยสังเกตจากการกระจายตัวของจุดบนแผนภาพความสัมพันธ์

4.3 ความแตกต่างของรูปภาพ (Differential Analysis)

ทดสอบโดยเปรียบเทียบความแตกต่างของภาพที่เข้ารหัสแล้วที่มีคีย์ที่แตกต่างกันเพียงเล็กน้อย [13] ซึ่งงานวิจัยนี้ใช้การคำนวณหาอัตราการเปลี่ยนแปลงของพิกเซล หรือที่เรียกว่า NPCR (Number pixel change rate) อัลกอริทึมในการคำนวณมีดังนี้

$$NPCR = \frac{\sum_{i,j} D(i,j)}{W \times H} \times 100\% \tag{14}$$

$$D(i,j) = \begin{cases} 1 & A(i,j) \neq B(i,j) \\ 0 & A(i,j) = B(i,j) \end{cases}$$

เมื่อ W และ H คือ ความกว้างและความยาวทั้งหมดของรูปภาพ

ตารางที่ 3: NPCR

รูปภาพ	คีย์ที่ 1	คีย์ที่ 2	NPCR
1	(14,37,56)	(10,37,56)	99.23 %
2	(86,14,70)	(86,15,70)	99.62 %
3	(2,56,32)	(2,56,33)	99.14 %

การทดสอบแบ่งเป็น 3 รูปแบบคือ 1.ค่าของกฎแตกต่างกัน 1 บิต 2.ค่า seedstate แตกต่างกัน 1 บิต 3.ค่า seedtime แตกต่างกัน 1 บิต

5. สรุป

งานวิจัยนี้ทำการปรับปรุงอัลกอริทึมในการเข้ารหัสและเพิ่มขั้นตอนในการเปลี่ยนภาพขาวดำให้เป็นภาพสีเทาเพื่อให้สามารถเข้ารหัสภาพขาวดำได้จากผลการทดลองแสดงให้เห็นว่างานวิจัยนี้สามารถปกปิดข้อมูลได้ดีกว่าวิธีเดิมมากและสามารถเข้ารหัสภาพขาวดำได้ด้วย

เอกสารอ้างอิง

- [1] วนิตา แก้วบุรณะประเสริฐ และ นันทิกา เบญจเทพานันท์, “การปรับปรุงวิธีการเข้ารหัสรูปภาพด้วยเซลล์ลูอาร์ออโตมาตาแบบพื้นฐาน” *8th Int. Joint Conf. on Computer Science and Software Engineering (JCSSE)*, 2011.
- [2] J. Jun, “Image encryption method based on elementary cellular automata” *SOUTHEASTCON*, vol. 9, 2009.
- [3] G. Allan, “The USC-SIPI Image Database:Version 5” [Online]. Available: <http://sipi.usc.edu/database/>, 2006.
- [4] J. Jun, “An image encryption based on elementary cellular automata” *Opt Laser Eng*, vol. 50, pp. 1836-1843, 2012.
- [5] J. Lang, “Image encryption based on the reality preserving multiple-parameter fractional Fourier transform and chaos permutation” *Opt Laser Eng*, vol. 50, pp. 929-937, 2012.
- [6] S.K. Rajput and Naveen K. Nishchal, “Image encryption and authentication verification using fractional non-conventional joint transform correlator” *Opt Laser Eng*, vol. 50,no. 10, pp. 1474-1483, 2012.
- [7] L. Zhang, X. Liao and X. Wang, “An image encryption approach based on chaotic maps” *Chaos Solitions and Fractals*, vol. 24, pp. 759-765, 2005.
- [8] N.K. Sreelaja and G.A. Vijayalakshmi Pai, “Stream cipher for binary image encryption using Ant Colony Optimization based key generation” *Applied Soft Computing*, vol. 12, pp. 2879-2895, 2012.
- [9] Wolfram, “Cryptography with cellular automata” *Crypto-89, Spriger*, vol. 218, pp. 429-432, 1986.
- [10] R. Chen and J. Lai, “Image security system using recursive cellular automata substitution” *Pattern Recognition*, vol. 40, no. 5, pp. 1621-1631, 2007.
- [11] R. Chen et al., “Image encryption/decryption system using 2-d cellular automata” *ISCE’06*, pp. 1-6, 2006.
- [12] Y. Li, L. Yuanxiang and Xia Xuewen, “Image encryption algorithm based on self-adaptive symmetrical-coupled toggle cellular automata” *2008 Congress on Image and Signal Processing*, vol. 3, pp. 32-36, 2008.
- [13] J. Ahmad and F. Ahmed, “Efficiency Analysis and Security Evaluation of Image Encryption Schemes” *IJVIPNS-IJENS*, vol. 12, pp. 18-31, 2012.

การสกัดคุณลักษณะภาพในการรู้จำใบหน้า Feature Extraction for Face Recognition

อุมพร สมทอง (Aumaporn Somthong)¹, ปิยวรรณ แสนปัญญา (Piyawan Sanpanyan)²,
สิรภัทร เชี่ยวชาญวัฒนา (Sirapat Chiewchanwattana)³ และ คำรณ สุนัตติ (Khamron Sunat)⁴

¹ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น

Sanpanya.piyawan@gmail.com, ouay.aumaporn@gmail.com, sunkra@kku.ac.th and khamron_sunat@yahoo.com

บทคัดย่อ

การรู้จำใบหน้าเป็นการยืนยัน พิสูจน์ตัวบุคคลด้วยภาพถ่าย ใบหน้าทั้งใบหน้าของบุคคล แต่ในปัจจุบันการรู้จำใบหน้าที่ยังมีประสิทธิภาพไม่ดีเท่าที่ควร งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อเพิ่มประสิทธิภาพของการรู้จำใบหน้าบุคคล โดยการใช้การสกัดคุณลักษณะภาพใบหน้าบุคคลแบบ PCA, LDA, 2DPCA และ 2DLDA เข้ามาใช้ในการเตรียมข้อมูลภาพ จากนั้นนำข้อมูลเข้าสู่ตัวแบบเครื่องจักรการเรียนรู้เอ็กซ์ตรีม (ELM) งานวิจัยนี้ใช้ภาพถ่ายใบหน้าที่ผู้วิจัยสร้างฐานข้อมูลขึ้นเองจำนวน 140 ภาพ เพื่อการเรียนรู้ของตัวแบบ และใช้ภาพจำนวน 60 ภาพ ในการทดสอบตัวแบบ เพื่อหาการสกัดคุณลักษณะภาพที่สามารถทำให้ตัวแบบเครื่องจักรการเรียนรู้เอ็กซ์ตรีม (ELM) สามารถรู้จำได้ดีที่สุด จากการเปรียบเทียบการทำงานของตัวแบบพบว่า การสกัดคุณลักษณะแบบ 2DLCA มีความถูกต้องของการสอนตัวแบบ 100 % และการทดสอบ 92.667 % มีประสิทธิภาพในการสกัดคุณลักษณะที่เหมาะสมกับเครื่องจักรการเรียนรู้เอ็กซ์ตรีม (ELM) มากที่สุด

คำสำคัญ: รู้จำใบหน้า การสกัดคุณลักษณะภาพ เครื่องจักรการเรียนรู้เอ็กซ์ตรีม PCA 2DPCA LDA 2DLDA

Abstract

Face recognition is used to verify the individual's identity by capturing his/her entire face image. The current approaches for face recognition, however, are found incompletely effective. Thus, this study aims to investigate a method to enhance the efficiency of

individual face recognition. The PCA-based, LDA-based, 2DPCA-based and 2DLDA-based methods are selected to use in extracting image features at the early stage of image data preparation. The processed data are transformed and classified further by using Extreme Learning Machine (ELM). The study employs a total self-developed database of 20 individual faces' images to better understand the model. All images are used to examine the model to propose an approach for image features extraction that allows ELM to recognize the faces precisely. From a comparison the performance of the 2DPCA model have teaching and testing is 100 % and 92.667 % respectively. The most efficient in extracting features that suit the extreme learning machine (ELM). The proposed approach for the most effective face recognition is also eventually expected.

Keyword: Face Detection, Image Feature Extraction, Evolutionary Extreme Learning Machine, PCA, 2DPCA, LDA, 2DLDA.

1. บทนำ

ในปัจจุบันการรู้จำภาพใบหน้าบุคคลเป็นที่นิยมกันอย่างแพร่หลาย และมีการนำมาประยุกต์ใช้ประโยชน์หลายด้าน เพราะการรู้จำภาพใบหน้าบุคคล ถือว่าเป็นการยืนยันตัวบุคคลเป็นการพิสูจน์เอกลักษณ์เฉพาะบุคคล โดยใช้คุณลักษณะเฉพาะทางกายภาพ (BIOMETRIC) [1] โดยส่วนใหญ่นำมาประยุกต์ใช้คืองานด้านการรักษาความปลอดภัย โดยเครื่องจักรเรียนรู้เอ็กซ์ตรีมจะทำงานโดยการเปรียบเทียบใบหน้าจากภาพถ่าย

ในฐานะข้อมูลใบหน้าที่มีอยู่ และเมื่อเครื่องจักรเรียนรู้เอ็กซ์ตรีมเปรียบเทียบใบหน้าเสร็จแล้วจะแสดงผลลัพธ์ใบหน้าบุคคลที่อยู่ที่มีใบหน้าเหมือนกับภาพต้นแบบที่นำมาเปรียบเทียบ [2] เครื่องมือเรียนรู้ภาพใบหน้าบุคคลนั้นได้รับความสนใจ และมีการพัฒนาปรับปรุงอย่างต่อเนื่อง จึงมีการพัฒนาอัลกอริทึมในการทำงานของเครื่องจักรเรียนรู้ออกมาหลากหลายรูปแบบ ซึ่งการพัฒนาอัลกอริทึมแต่ละยุคก็แตกต่างกัน เนื่องจากเทคโนโลยีอุปกรณ์ และปัจจัยด้านองค์ความรู้ที่ต่างกัน จึงทำให้ต้องออกแบบอัลกอริทึมใหม่ให้เหมาะสมกับยุคสมัย และเทคโนโลยีที่สูงขึ้น จึงทำให้เครื่องจักรเรียนรู้ภาพใบหน้าบุคคลมีความแม่นยำและน่าเชื่อถือมากขึ้น

ด้วยเหตุที่กล่าวมาข้างต้นนี้ จึงเกิดแนวคิดในการทำเครื่องมือรู้จำภาพใบหน้าบุคคล โดยออกแบบเครื่องจักรเรียนรู้สามารถตรวจจับใบหน้าบุคคล (FACE DETECTION) และ การรู้จำใบหน้าบุคคล (FACE RECOGNITION) ได้ในเครื่องจักรเดียว โดยภาพที่นำมาวิเคราะห์จะใช้ภาพถ่ายดิจิทัล และภาพจากกล้องถ่ายภาพวิดีโอ แล้วนำภาพที่ได้มาผ่าน โปรแกรมช่วยในการวิเคราะห์ภาพ โดยใช้เทคนิคการสกัดคุณลักษณะภาพดิจิทัล PCA, 2DPCA, 2DLDA และ LDA เข้ามาใช้ในการเตรียมข้อมูลเพื่อให้เครื่องจักรเรียนรู้ทำงานได้ถูกต้องแม่นยำมากขึ้น แล้วนำภาพที่ได้จากการสกัดคุณลักษณะภาพ มาผ่านข่ายงานเครื่องจักรเรียนรู้แบบเอ็กซ์ตรีม (ELM) แล้วนำผลลัพธ์ที่ได้จากการรู้จำใบหน้ามาเพื่อวิเคราะห์หาความถูกต้อง และแม่นยำของการสกัดคุณลักษณะภาพแต่ละเทคนิคว่าเทคนิคใดสามารถรู้จำใบหน้าได้ถูกต้องและแม่นยำมากกว่า เพื่อให้ได้การสกัดคุณลักษณะภาพที่เหมาะสมที่สุดในการสร้างเครื่องจักรเรียนรู้จำภาพใบหน้าบุคคลที่แม่นยำและดีที่สุด

2. งานวิจัยที่เกี่ยวข้อง

2.1 เครื่องจักรเรียนรู้เอ็กซ์ตรีม (ELM)

Extreme learning machine (ELM) [3] เป็นขั้นตอนวิธีการที่ได้รับความนิยมจากนักวิจัยซึ่ง ELM นี้จะใช้งาน ได้กับสถาปัตยกรรมโครงข่ายประสาทเทียมแบบป้อนไปข้างหน้าที่มีโหนดชั้นซ่อนเพียงชั้นเดียว (Single-hidden layer feed-forward

networks : SLFNs) และที่สำคัญ ELM ไม่ต้องมีการวนปรับค่าน้ำหนักในโหนดชั้นซ่อนของ SLFNs และเมื่อเปรียบเทียบเทคนิคการคำนวณแบบดั้งเดิมแล้ว ELM มีความเร็วสูงกว่า [4]

2.2 การวิเคราะห์ส่วนประกอบสำคัญ (PCA)

เป็นวิธีการที่ใช้ในทฤษฎีไอเกนเฟส มาใช้แทนข้อมูลเวกเตอร์ (Eigenvector) โดยขั้นตอนการทำงานสำคัญคือ ทำการแปลงโครงสร้างเมตริกซ์ข้อมูลเป็นเวกเตอร์แล้ว คำนวณหาไอเกนเวกเตอร์ที่สอดคล้องกันกับค่าไอเกน แล้วนำไปคำนวณหาลักษณะเด่นของภาพเพื่อนำไปใช้หาการรู้จำ [5]

2.3 การวิเคราะห์การแบ่งแยกเชิงเส้น (LDA)

คือการแบ่งประเภทวัตถุ เป็นหนึ่งในฐานสองกลุ่มหรือมากกว่าสองกลุ่ม เป็นชุดของคุณลักษณะซึ่งเป็นรายละเอียดของวัตถุ โดยทั่วไปเรากำหนดวัตถุเป็นหนึ่งจำนวนฐานกลุ่มที่ได้กำหนดไว้ล่วงหน้า การสังเกตที่ทำบนวัตถุสามารถใช้รูปแบบการจำแนกเชิงเส้น (LDA) ซึ่งแยกได้เป็นเส้นตรงแสดงให้เห็นว่ากลุ่มที่สามารถแยกจากกัน โดย การรวมกันเชิงเส้นของคุณลักษณะซึ่งเป็นรายละเอียดของวัตถุ [6]

2.4 การวิเคราะห์แบ่งแยกเชิงเส้นสองมิติ (2DLDA)

วิธีการ 2DLDA เป็นวิธีการใหม่ที่ประยุกต์มาจากวิธี FLD และ 2DPCA โดยสามารถนำเมทริกซ์ภาพมากระจายในกลุ่มและกระจายระหว่างกลุ่ม แล้วนำมาคำนวณหาค่าไอเกนเวกเตอร์ที่สมนัยกับค่าไอเกน ที่เรียงจากมากไปหาน้อย แล้วนำภาพใบหน้ามาฉายลงบนไอเกนเวกเตอร์ จะได้คุณลักษณะเด่นที่มีจำนวนของคอลัมน์ภาพน้อยลง

3. วิธีการดำเนินงาน

ในส่วนนี้เป็นการเสนอวิธีการสกัดคุณลักษณะในการรู้จำใบหน้าด้วยเทคนิคการสกัดภาพแบบ PCA, 2DPCA, LDA และ 2DLDA โดยใช้ตัวแบบในการรู้จำแบบเอ็กซ์ตรีมในการรู้จำใบหน้า เพื่อหาการสกัดคุณลักษณะภาพที่เหมาะสมที่สุดที่ช่วยให้การรู้จำใบหน้าด้วยเครื่องจักรเรียนรู้เอ็กซ์ตรีมสามารถทำงานได้ประสิทธิภาพมากที่สุด ในขั้นตอนแรกผู้วิจัยได้จัดเตรียมฐานข้อมูลภาพใบหน้าบุคคลของตนเอง โดยใช้ข้อมูลจาก CSKKU Database ซึ่งภายในฐานข้อมูลประกอบไป

ด้วยภาพใบหน้าบุคคล 20 คน เป็นเพศชาย 10 คน และเป็นเพศหญิง 10 คนซึ่งทั้ง 20 คนเป็นนักศึกษาในภาควิชาวิทยาการคอมพิวเตอร์ทั้งหมด และผู้วิจัยใช้ภาพจากฐานข้อมูล ESSEX Database ซึ่งเป็นฐานข้อมูลสำหรับการทำ Biometric โดยเฉพาะมาใช้ในการทดสอบเครื่องมือรู่วมด้วย โดยฐานข้อมูล ESSEX Database ประกอบไปด้วยภาพใบหน้าบุคคลจำนวน 200 ภาพ แบ่งเป็นเพศชาย 16 คน และเพศหญิง 4 คน จากนั้นผู้วิจัยทำการปรับขนาดเป็น 50 x 60 พิกเซล และแบ่งเป็นชุดข้อมูลเท่ากัน โดยใช้อัตราส่วน 70:30 ซึ่งแบ่งเป็นภาพสำหรับสอนตัวแบบจำนวน 140 ภาพ และ 60 ที่เหลือสำหรับทดสอบตัวแบบ โดยฐานข้อมูลทั้งสองที่ผู้วิจัยใช้มีการปรับขนาดภาพ และมีจำนวนภาพที่ใช้เท่ากัน จากนั้นนำภาพที่เตรียมเรียบร้อยแล้วนำไปสกัดคุณลักษณะภาพทั้ง 4 แบบ

โดยแบบที่ 1 คือแบบ การวิเคราะห์องค์ประกอบหลัก (PCA) โดยนำภาพทั้ง 2 ชุดข้อมูลมาแปลงเป็นภาพระดับสีเทาแล้วทำการคำนวณหาหลักเกณฑ์เด่นของภาพเพื่อนำไปใช้การรู้จำ โดยกำหนดให้ภาพที่ใช้ฝึกสอน P และภาพมีขนาด $M \times N$ คือ x_1, x_2, \dots, x_p และแปลงเมตริกซ์ภาพให้เรียงต่อกันเป็นเวกเตอร์ขนาด $1 \times d$, $d = m \times n$ จะใช้ชุดข้อมูลฝึกสอนดังนี้ $\mu = \{\mu_1, \mu_2, \dots, \mu_N\}$ จากนั้นนำมาคำนวณหาค่าเฉลี่ยของภาพใบหน้าเพื่อหาค่ากลางข้อมูลใหม่ จากนั้นคำนวณหาค่าความแปรปรวนร่วมของข้อมูลเมตริกซ์โควาเรียนซ์ เมื่อแทนค่าเฉลี่ยโดยที่ C เป็นเมตริกซ์ความแปรปรวนร่วม μ และนำค่า C ไปคำนวณหาไอเกนเวกเตอร์ที่สอดคล้องกับค่าไอเกนสอดคล้องกับค่าไอเกนจากมากไปน้อย แล้วคัดเอาเฉพาะค่าที่ไม่เท่ากับศูนย์ จากนั้นทำการสกัดเอาหลักเกณฑ์เด่นของภาพเพื่อนำไปใช้หาการรู้จำจาก

$$Y_k = U_d^T (\mu_k - \mu); k=1, 2, \dots, P \quad (1)$$

เมื่อแทน Y_k = ลักษณะเด่นภาพที่ k

d = ค่ามิติที่เลือกระหว่าง 1 ถึง $P-1$

และ k = จำนวนภาพ P ภาพ [7]

การสกัดคุณลักษณะภาพแบบที่ 2 วิธี 2DPCA เป็นเทคนิคที่สามารถนำเอาภาพสองมิติมาคำนวณค่าได้โดยตรง โดยใช้หลักการคำนวณคือเริ่มจากการนำภาพจากฐานข้อมูลชุดที่ 1

และชุดที่ 2 มาแปลงภาพเป็น Gray Scale ทั้งหมด นำภาพมาเรียงเป็นเมตริกซ์ภาพ โดยกำหนดให้ X ประกอบด้วยเวกเตอร์ขนาด n มิติ และ A คือภาพในฐานข้อมูลที่มีขนาด 50×60 พิกเซลจากนั้นนำภาพ A มาฉายภาพลงบนแกน X ตามหลักการแปลงเชิงเส้น จะได้ $Y = AX$ โดยที่ Y คือ พีเจอร์เวกเตอร์ของภาพ A นำภาพจากข้อมูลชุดการสอนตัวแบบทั้งหมด 140 ภาพ ซึ่งเป็นเมตริกซ์ขนาด 50×60 โดยหาค่าเฉลี่ยของแต่ละภาพแล้วนำค่าเฉลี่ยไปหาค่าความแปรปรวนร่วมของภาพ คำนวณหาค่าไอเกนเวกเตอร์ที่สมนัยกับค่าไอเกน จากเมตริกซ์ความแปรปรวนร่วม แล้วเรียงลำดับค่าไอเกนเวกเตอร์จากมากไปหาน้อย แล้วเลือกแกนการฉายภาพกลับเพื่อดูมิติของภาพจะนำไปเลือกใช้ในการรู้จำต่อไป

ขั้นตอนการสกัดคุณลักษณะภาพแบบที่ 3 2DLDA มีหลักการคำนวณดังนี้ เริ่มต้นด้วยแปลงภาพทั้งหมดเป็น Gray Scale จากนั้นคำนวณค่าเฉลี่ยทั้งหมดเพื่อหาค่ากลางของชุดข้อมูลคำนวณเมตริกซ์การกระจายของภาพภายในกลุ่มของแต่ละกลุ่ม ซึ่งการคำนวณจะใช้การหาผลรวมของ เมตริกซ์ความแปรปรวนร่วมในแต่ละกลุ่ม เมื่อมีภาพทั้งหมด L กลุ่ม คำนวณเมตริกซ์การกระจายของภาพระหว่างกลุ่ม ซึ่งการคำนวณจะใช้การหาผลรวมของเมตริกซ์ความแปรปรวนร่วมเฉลี่ยระหว่างกลุ่ม คำนวณค่าไอเกนเวกเตอร์และค่าไอเกนจากเมตริกซ์การกระจายภายในกลุ่มและเมตริกซ์การกระจายระหว่างกลุ่ม จากนั้นเรียงค่าไอเกนเวกเตอร์ให้สอดคล้องกับค่าไอเกนจากมากไปน้อย แล้วเลือกแกนการฉาย 1 ถึง d โดยที่ $d \leq n$ คำนวณค่าคุณลักษณะของภาพจากสมการ

$$Y = AX \quad (2)$$

ซึ่ง A = เมตริกซ์ภาพที่มีขนาด $m \times n$ มิติ

X = เวกเตอร์ขนาด n มิติ

และ Y = เวกเตอร์คุณลักษณะของเมตริกซ์ภาพ A

การสกัดคุณลักษณะภาพแบบที่ 4 LDA พยายามที่จะลดมิติข้อมูลพร้อมทั้งรักษาข้อมูลการวิเคราะห์ที่เท่าที่จะเป็นไปได้ ขั้นตอนการสกัดคุณลักษณะภาพ เริ่มขั้นตอนจากแปลงภาพจาก

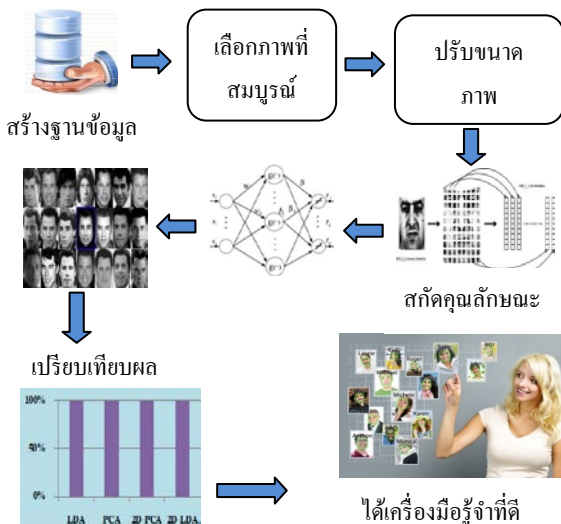
ฐานข้อมูลเป็น Gray Scale จากนั้นแปลงภาพแต่ละภาพเป็นเวกเตอร์ขนาด $1 \times N$ มิติ จากนั้นหาค่าเฉลี่ยของข้อมูลทั้งหมด จากนั้นคำนวณหาการกระจายภายในกลุ่ม S_w จากนั้นคำนวณเมตริกซ์การกระจายระหว่างกลุ่ม S_b จากนั้นคำนวณหาค่าไอเกนเวกเตอร์ และค่าไอเกนของ S_w และ S_b เมื่อได้ค่าไอเกนแล้ว นำไอเกนเวกเตอร์ที่สมนัยกับค่าไอเกนจากมากไปหาน้อย จากนั้นคำนวณหาค่าคุณลักษณะด้วยสมการ

$$Y = W^T t \tag{3}$$

ซึ่ง Y = เวกเตอร์คุณลักษณะของเมตริกซ์ภาพ W

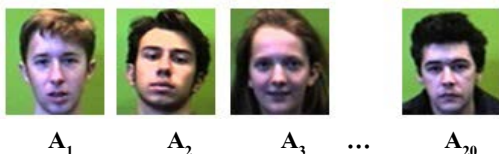
W = ค่าไอเกน

เมื่อกำหนดการสกัดคุณลักษณะของภาพใบหน้าบุคคลครบทั้ง 4 แบบแล้ว จากนั้นนำข้อมูลเข้าสู่ตัวแบบเอ็กซ์ตรีมเพื่อทำการรู้จำใบหน้าแล้วดูผลลัพธ์ และทำการประเมินผล



ภาพที่ 1: แผนภาพขั้นตอนการรู้จำใบหน้า

ภาพที่ 2: ตัวอย่างภาพจากฐานข้อมูล CSKKU Database



ภาพที่ 3: ตัวอย่างภาพจากฐานข้อมูล ESSEX Database

4. ผลการดำเนินการ

ผลการดำเนินงานของงานวิจัยนี้เป็นการเปรียบเทียบประสิทธิภาพการสกัดคุณลักษณะภาพแบบ PCA, 2DPCA, LDA และ 2DLDA ในการรู้จำใบหน้าด้วยเครื่องจักรการเรียนรู้เอ็กซ์ตรีม ทำการประมวลโดยโปรแกรม MATLAB สำหรับข้อมูลที่ใช้ทั้งหมด 200 ภาพ ซึ่งแบ่งเป็นข้อมูลภาพชุดที่ 2 มีภาพจำนวน 140 ภาพใช้สำหรับสอนตัวแบบ และชุดที่ 2 มีภาพจำนวน 60 ภาพใช้ในการทดสอบตัวแบบ ซึ่งการทำเครื่องมือจากการทดสอบเปรียบเทียบวิธีการสกัดคุณลักษณะภาพแบบ PCA, LDA, 2DPCA และ 2DLDA การวัดประสิทธิภาพในแต่ละรอบจะได้ค่าความถูกต้องของการสอน และทดสอบตัวแบบที่แตกต่างกัน ซึ่งงานวิจัยนี้ได้ทดสอบความถูกต้องจำนวน 50 รอบ โดยใช้ภาพชุดข้อมูลภาพที่มีขนาด 50 x 60 พิกเซล ซึ่งเป็นภาพจากฐานข้อมูลของผู้วิจัยสร้างขึ้นเอง CSKKU Database และใช้ภาพจากฐานข้อมูล ESSEX Database ที่เป็นฐานข้อมูลใบหน้าบุคคลแบบสาธารณะเพื่อการวิจัยของ University of Essex, UK [8] มาใช้ในการประมวลผลการรู้จำใบหน้าบุคคลรวมด้วย เนื่องจากผู้วิจัยต้องการความหลากหลายของผลการวิจัย และต้องการทดสอบตัวแบบการเรียนรู้ที่ผู้วิจัยสร้างขึ้นนี้ เหมาะสมกับฐานข้อมูลอื่นๆ ด้วยหรือไม่ หลังจากประมวลผลภาพจากทั้ง 2 ฐานข้อมูลแล้ว ผู้วิจัยจะเปรียบเทียบค่ามิติ ของแต่ละการสกัดคุณลักษณะว่าสามารถสกัดคุณลักษณะได้มากน้อยแตกต่างกันอย่างไร และผู้วิจัยจะนำค่าความถูกต้องมาเฉลี่ยหาค่าร้อยละของความถูกต้อง ที่แสดงในตารางที่ 2 - 5 โดยในส่วนผลลัพธ์ที่ถูกทำเป็นช่องที่บหมายถึงผลลัพธ์ที่แสดงค่าร้อยละของความถูกต้องในการรู้จำ และทดสอบข้อมูลสูงสุดของแต่ละตาราง หรือแสดงถึงค่าสูงสุดที่ต้องการเปรียบเทียบในตารางโดยในการวิจัยครั้งนี้ผู้วิจัยยังแสดงผลของค่ามิติ ที่ได้จากการสกัดคุณลักษณะแต่ละคุณลักษณะ เพื่อแสดงค่ามิติ ที่การสกัดคุณลักษณะแต่ละแบบที่สามารถคำนวณหาได้จากการสกัดคุณลักษณะภาพขนาด 50 x 60 พิกเซล โดยใช้ภาพจากฐานข้อมูลของผู้วิจัย และใช้ภาพจากฐานข้อมูล ESSEX Database ได้ค่ามิติ ไอเกนเวกเตอร์ดังนี้

ตารางที่ 1: เปรียบเทียบขนาดมิติภาพขนาด 50 x 60 พิกเซล ภาพจาก

ฐานข้อมูล CSKKU Database และ ESSEX Database

Database	PCA	LDA	2DPCA	2DLDA
CSKKU Database	3000, 138	3000, 19	60, 50	60, 50
ESSEX Database	3000, 138	3000, 19	60, 50	60, 50

จากตารางที่ 1 จะพบว่าการสกัดคุณลักษณะของภาพจากฐานข้อมูล CSKKU Database และภาพจากฐานข้อมูล ESSEX Database สามารถสกัดคุณลักษณะภาพได้จำนวนเท่ากันทุกการสกัดคุณลักษณะ จากตารางด้านบนพบว่าการสกัดคุณลักษณะแบบ PCA สามารถหาค่ามิติได้มากที่สุดรองลงมาเป็น การสกัดคุณลักษณะแบบ 2DPCA และ 2DLDA ได้ค่ามิติเท่ากัน และ LDA สกัดคุณลักษณะภาพได้น้อยที่สุด

ตารางที่ 2: การเปรียบเทียบประสิทธิภาพการสกัดคุณลักษณะภาพระหว่าง PCA และ LDA โหนดคงที่ 1000 โหนด ขนาดภาพ 50 x 60 พิกเซล ภาพจาก CSKKU Database

มิติ	PCA+ELM			LDA+ELM		
	Train	Test	Time	Train	Test	Time
50	100	90.333	0.310	-	-	-
19	100	86.333	0.240	100	22.467	235.829
10	100	55.667	0.227	100	13.667	235.040
เฉลี่ย	100	77.444	0.259	100	18.067	235.434

ตารางที่ 3: การเปรียบเทียบประสิทธิภาพการสกัดคุณลักษณะภาพระหว่าง 2DPCA และ 2DLDA โหนดคงที่ 1000 โหนด ขนาดภาพ 50 x 60 พิกเซล ภาพจาก CSKKU Database

มิติ	2DPCA+ELM			2DLDA+ELM		
	Train	Test	Time	Train	Test	Time
50	100	89.333	0.380	100	89.667	0.394
19	100	93.679	0.266	100	93.667	0.281
10	100	94.333	0.235	100	94.667	0.240
เฉลี่ย	100	92.449	0.293	100	92.667	0.305

จากตารางที่ 2 จะพบช่องมิติที่ 50 ของ LDA+ELM ไม่มีผลการทดลอง เนื่องจากการสกัดคุณลักษณะแบบ LDA สามารถคำนวณค่าไอเกนเวกเตอร์ได้จำนวน 19 มิติ ในส่วนตารางที่ 2 - 3 เป็นการเปรียบเทียบประสิทธิภาพของเครื่องมือรู้จำที่ใช้ภาพจากฐานข้อมูลของผู้วิจัยในการทดสอบเครื่องมือรู้จำเปรียบเทียบประสิทธิภาพ โดยการกำหนดโหนดจำนวน 1000 โหนด ให้คงที่ตลอดการทดสอบเครื่องมือรู้จำ และทำการปรับค่ามิติ ตั้งแต่ 10, 19 และ 50 มิติ ตามลำดับ พบว่าการสกัดคุณลักษณะภาพแบบ 2DLDA สามารถรู้จำได้ดีที่สุด และค่าประสิทธิภาพของการทดสอบตัวแบบ และเวลาการคำนวณผลได้ดีที่สุด

ตารางที่ 4: การเปรียบเทียบประสิทธิภาพการสกัดคุณลักษณะภาพระหว่าง PCA และ LDA โหนดคงที่ 1000 โหนด ขนาดภาพ 50 x 60 พิกเซล ภาพจากฐานข้อมูล ESSEX Database

มิติ	PCA+ELM			LDA+ELM		
	Train	Test	Time	Train	Test	Time
50	100	100	0.313	-	-	-
19	100	97.667	0.230	100	24	236.222
10	100	50.667	0.233	100	21.867	284.706
เฉลี่ย	100	82.778	0.259	100	22.934	260.464

จากตารางที่ 4 ช่องใดเมนชั้นที่ 50 ของ LDA+ELM ไม่มีผลการทดลองเนื่องจากไอเกนเวกเตอร์ของตัวแบบ LDA มีค่าสูงสุดเท่ากับ 19 มิติ

ตารางที่ 4 - 5 เป็นการปรับค่าจำนวนโหนดที่ใช้ในการคำนวณเป็น 1000 โหนด โดยกำหนดให้จำนวนโหนดคงที่ตลอดการประมวลผล ผู้วิจัยใช้จำนวนรอบในการรู้จำทั้งหมด 50 รอบ โดยใช้ฐานข้อมูลภาพจาก ESSEX Database และกำหนดค่ามิติของภาพขึ้นมาโดยแต่ละค่า จะปรับใหม่ทุกรอบของการทดสอบ เพื่อตรวจสอบประสิทธิภาพในการรู้จำและการทดสอบตัวแบบว่าเมื่อมีการเปลี่ยนแปลงจำนวนโหนดและเปลี่ยนแปลงค่ามิติแล้วจะมีผลเป็นอย่างไร

ตารางที่ 5: การเปรียบเทียบประสิทธิภาพการสกัดคุณลักษณะภาพระหว่าง 2DPCA และ 2DLDA โหนดคงที่ 1000 โหนด ขนาดภาพ 50 x 60 พิกเซล ภาพจากฐานข้อมูล ESSEX Database

ชื่อ มิติ	2DPCA+ELM			2DLDA+ELM		
	Train	Test	time	Train	Test	time
50	100	99.667	0.372	100	99.667	0.381
19	100	100	0.257	100	100	0.274
10	100	99.333	0.213	100	100	0.231
เฉลี่ย	100	99.667	0.281	100	99.889	0.295

จากตารางที่ 4 - 5 พบว่าการสกัดคุณลักษณะแบบ 2DLDA จะมีประสิทธิภาพในการรู้จำได้ดีกว่าการสร้างเครื่องมือรู้จำด้วยการสกัดคุณลักษณะภาพแบบ PCA, LDA และ 2DPCA โดยใช้ภาพจาก ESSEX Database

5. สรุปผลการวิจัย

ในการวิจัยเพื่อหาการสกัดคุณลักษณะภาพที่เหมาะสมกับการรู้จำใบหน้าด้วยเครื่องจักรการเรียนรู้แบบเอ็กซ์ตรีม ผู้วิจัยตัดสินใจเลือกผลการวัดประสิทธิภาพการวิจัยจากตารางที่ 3 ที่มีการเปรียบเทียบผลการวัดประสิทธิภาพ จากการรู้จำภาพจากฐานข้อมูล CSKKU Database จากการทดลองจะแสดงให้เห็นว่าการสกัดคุณลักษณะภาพด้วยวิธี 2DLCA ทำให้การรู้จำใบหน้าด้วยเครื่องจักรการเรียนรู้แบบเอ็กซ์ตรีม โดยใช้มิติเท่ากับ 10 มิติ จะทำให้ค่าความถูกต้องโดยเฉลี่ยของการสอนตัวแบบ และการทดสอบตัวแบบมีค่ามากกว่าทุกวิธีในการทดสอบ และในส่วนของตารางที่ 6 จะเปรียบเทียบให้เห็นว่าเทคนิค 2DLDA มีค่าเฉลี่ยของประสิทธิภาพมากกว่าทุกวิธีที่ใช้ในการทดสอบ มีค่าประสิทธิภาพการสอนตัวแบบมีความถูกต้องเท่ากับร้อยละ 100 และมีความถูกต้องของการทดสอบตัวแบบเท่ากับร้อยละ 92.667 และจากตารางที่ 6 พบว่าการทดสอบตัวแบบในการรู้จำด้วยข้อมูลภาพจาก ESSEX Database ก็พบว่าประสิทธิภาพการสกัดคุณลักษณะภาพแบบ 2DLDA มีประสิทธิภาพมากที่สุดด้วยเช่นเดียวกับการใช้ฐานข้อมูลของผู้วิจัย

ตารางที่ 6: การเปรียบเทียบประสิทธิภาพโดยเฉลี่ยของการสกัดคุณลักษณะ แบบ PCA, 2DPCA, LDA และ 2DLDA จำนวนโหนด 1000 โหนด ภาพขนาด 50 x 60 มิติ

Feature Extraction	CSKKU Database		ESSEX Database	
	Test	Time	Test	Time
PCA	77.444	0.259	82.778	0.259
LDA	18.067	235.434	22.934	260.464
2DPCA	92.449	0.293	99.667	0.281
2DLDA	92.667	0.305	99.889	0.295

เอกสารอ้างอิง

- [1] I-TOP Technology Company.(2550) ไบโอมเทรคซ์ (Biometric) คืออะไร. สืบค้นเมื่อ 21 พฤษภาคม 2556, เข้าถึงได้จาก http://www.hitop.co.th/Biometric_detail.htm
- [2] BANTRONIX. (2554). การรู้จำใบหน้า. สืบค้นเมื่อ 21 พฤษภาคม 2556, เข้าถึงได้จาก : <http://www.bantronix.com/2011/10/face-recognition.html>
- [3] Huang, GB., Wang, DH., and Lan Y. (2011). Extreme learning machines: a survey. *International Journal of Machine Learning and Cybernetics*, 2011 Vol.2, 107-122 doi: 10.1007/s13042-011-0019-y
- [4] ธนาวุฒิ ประกอบผล. (2552). “โครงข่ายประสาทเทียม”. วารสาร มจร.วิชาการ, ปีที่ 12(24), 73-87.
- [5] Turk, M., and Pentland, A. (1991). *Eigenfaces for recognition*. *Journal of Cognitive Neuroscience*. 3:Page 71-86, Received July 15,2013 from <http://www.sciencedirect.com>
- [6] L10: Linear discriminants analysis. Received July 23, 2013, from http://research.cs.tamu.edu/prism/lectures/pr/pr_110.pdf
- [7] Kadi Tecknomo. *LINEAR DISCRIMINANT ANALYSIS (LDA)*. Received July 23, 2013, from <http://people.revoledu.com/kardi/tutorial/LDA/LDA.html>.
- [8] Libor Spacek. (2008). *Description of the Collection of Facial Images*. Received February 26, 2014. form <http://cswww.essex.ac.uk/mv/allfaces/index.html>

การศึกษาการใช้กล้องซีซีดีในการตรวจวัดปริมาณไลโคปีนในมะเขือเทศ ด้วยโครงข่ายประสาทเทียม

The Study of Using CCD Camera to Measure the Lycopene in Tomato by Colors with Artificial Neuron Networks

รพีพงศ์ รัตนวรหิรัญกุล (Rapeepong Rattanawaorahirunkul)¹ สราวุฒิ สุพรรณราช (Sarawut Supannarach)²
และจิตสรายุ สীগูกา (Jitsaran Seekuka)³

^{1,2}สาขาวิชาวิศวกรรมไฟฟ้าและคอมพิวเตอร์ คณะวิทยาศาสตร์และวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์
วิทยาเขตเฉลิมพระเกียรติ จังหวัดสกลนคร

³ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตบางเขน
Rapeepong EE@yahoo.com, apor ku@hotmail.com, ๕5414550301@ku.ac.th

บทคัดย่อ

ไลโคปีนเป็นสารสำคัญที่พบในมะเขือเทศ ซึ่งมีส่วนช่วยในการต้านอนุมูลอิสระและสามารถป้องกันการเกิดมะเร็งต่อมลูกหมากได้ ปัจจุบันมีการวิเคราะห์ปริมาณสารไลโคปีนในมะเขือเทศ โดยใช้วิธีโครมาโทกราฟีและวิธีสเปกโตรโฟโตเมตรี-อัลตราไวโอเลตซึ่งมีขั้นตอนที่ซับซ้อนและต้องใช้สารเคมีเฉพาะในการวิเคราะห์ ส่งผลให้ต้นทุนในการวิเคราะห์สูงงานวิจัยนี้จึงมีวัตถุประสงค์เพื่อนำเสนอเทคนิคการวัดปริมาณสารไลโคปีนในมะเขือเทศ โดยใช้ภาพถ่ายจากกล้องซีซีดี โดยนำตัวอย่างมะเขือเทศมาถ่ายภาพ และวัดค่าสีจากภาพถ่ายในระบบ สีแดง สีเขียว และสีน้ำเงินด้วยโครงข่ายประสาทเทียม แล้วนำมาเปรียบเทียบกับวิธีการวิเคราะห์ปริมาณสารไลโคปีนด้วยวิธีสเปกโตรโฟโตเมตรีอัลตราไวโอเลต จากผลการวิเคราะห์พบว่าปริมาณสารไลโคปีนที่วัดได้โดยใช้โครงข่ายประสาทเทียมอ้างอิงจากภาพถ่ายในระบบสีแดง สีเขียว และสีน้ำเงินของมะเขือเทศสดที่ได้จากกล้องซีซีดี เทียบกับการวิเคราะห์ด้วยวิธีสเปกโตรโฟโตเมตรีอัลตราไวโอเลต พบว่า Multi Layer Feed Forward Back Propagation Neuron Network มีค่ากลางของความคลาดเคลื่อนสัมบูรณ์ 0.6520 ค่ากลางของความคลาดเคลื่อนกำลังสอง 0.3341 และค่ากลางของเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์ 0.4762 ตามลำดับ งานวิจัยนี้แสดงให้เห็นว่าการใช้เซ็นเซอร์ตรวจวัดปริมาณไลโคปีนในมะเขือเทศมีความเชื่อถือได้สูงและมีความไม่แตกต่างกันอย่างมีนัยสำคัญ

เมื่อเทียบกับการวัด โดยวิธีสเปกโตรโฟโตเมตรีอัลตราไวโอเลตที่มีค่าใช้จ่ายและใช้เวลามากกว่า

คำสำคัญ: สารไลโคปีน ระบบสี โครงข่ายประสาทเทียม สเปกโตรโฟโตเมตรีอัลตราไวโอเลต

Abstract

The Lycopene is an important substance found in Tomato which can help in the antioxidant and can prevent prostate cancer. At present, the conventional methods to analyze the Lycopene amount in Tomato are Chromatography and UV-Spectrophotometer. However, these two methods require several complicated processes and specific chemical substances which elevated the analysis cost.

This research aims to present the use of images from CCD camera for measuring the amount of Lycopene in Tomato. The measured data of fresh Tomato sample from the color sensor in RGB color system by neuron networks was compared with amount of Lycopene obtained from the conventional method of UV-Spectrophotometry. In addition, the analysis result from the color sensor implemented in this research and the conventional UV-Spectrophotometer found that Multi Layer Feed Forward Back Propagation Neuron Network

Have Mean Absolute Error (MAE) 0.6520, Mean Square Error (MSE) 0.3341, and Mean Absolute Percentage Error (MAPE) 0.4762 respectively. This research shows that the color sensor can be used to analyze the amount of Lycopene in Tomato with high reliability that not significantly different from the conventional UV-Spectrophotometer method which take more cost and time to do the Lycopene measurement analysis.

Keywords: Lycopene, Color System, Neuron Network, UV Spectrophotometer.

1. บทนำ

มะเขือเทศเป็นพืชสวนครัวชนิดหนึ่งที่กำลังได้รับความนิยมอย่างแพร่หลายเพราะมีประโยชน์ ทั้งในด้านการด้านสารอนุมูลอิสระ และยังมีวิตามินต่างๆอีกมากมาย ซึ่งสารสำคัญที่พบในมะเขือเทศ คือ สารไลโคปีน (Lycopene) และวิธีวิเคราะห์ปริมาณสารไลโคปีนในปัจจุบันนิยมใช้เทคนิคการวิเคราะห์ด้วยวิธีโครมาโตกราฟี (Chromatography) และวิธีสเปกโตรโฟโตเมตรีอัลตราไวโอเลต (UV-Spectrophotometry) ซึ่งทั้ง 2 วิธีนี้จะต้องใช้เวลานานมีขั้นตอนที่ซับซ้อนต้องใช้สารเคมีในการวิเคราะห์จึงทำให้ต้นทุนในการวิเคราะห์สูง [1] จากปัญหาข้างต้นงานวิจัยนี้มีวัตถุประสงค์ในการพัฒนาวิธีวิเคราะห์ปริมาณสารไลโคปีนให้มีความสะดวก รวดเร็ว และได้ค่าที่ถูกต้อง งานวิจัยนี้จึงได้ประยุกต์ใช้การประมวลผลสัญญาณภาพ (Image Processing) บนสัญญาณภาพ 2 มิติ จากการถ่ายภาพสีด้วยกล้อง CCD โดยนำภาพดังกล่าวมาวิเคราะห์ข้อมูลสี ด้วยการใช้เทคนิควิธีแยกข้อมูลแต่ละจุดสีในภาพถ่ายออกเป็นข้อมูลภายในปริภูมิ 3 มิติ หรือแยกข้อมูลออกเป็นค่าในแกนสีแดง (R) สีเขียว (G) และสีน้ำเงิน (B) เพื่อนำค่าในแกนสีทั้งสามแกนมาเข้าสู่กระบวนการของโครงข่ายประสาทเทียม การวิเคราะห์ผลการทดลองทำได้โดยนำตัวอย่างมะเขือเทศมาตรฐานวัดวิเคราะห์หาสีในแต่ละแกนสี จากภาพถ่ายที่ได้จากกล้องซีซีดี โดยนำข้อมูลดังกล่าวเข้าสู่กระบวนการประมวลผลภาพเพื่อให้ทราบอัตราการเปลี่ยนแปลงของค่าสีในแต่ละแกนสีที่มีความสัมพันธ์กับความเข้มสีของมะเขือเทศสด จากนั้นนำตัวอย่างมะเขือเทศสดดังกล่าวไปทำการวิเคราะห์หาปริมาณสารไลโคปีนด้วยเทคนิควิธี สเปกโตรโฟโตเมตรีอัลตราไวโอเลต

งานวิจัยนี้ทำการทดสอบค่าความถูกต้องของโครงข่ายประสาทเทียมโดยทำการเปรียบเทียบกับปริมาณสารไลโคปีนที่วัดได้ด้วยเทคนิควิธีสเปกโตรโฟโตเมตรีอัลตราไวโอเลต จากนั้นนำข้อมูลที่ได้มาวิเคราะห์ทางสถิติเพื่อหาผลสรุปต่อไป

2. ทฤษฎีที่เกี่ยวข้อง

2.1 มาตรฐานของสี

จากข้อมูลมาตรฐานของสีที่ใช้อยู่ในปัจจุบันมีอยู่หลายระบบด้วยกันทั้งนี้จะขึ้นอยู่กับนำไปใช้แต่โดยทั่วไปแล้วทุกมาตรฐานจะมีแนวคิดเดียวกันคือการแทนจุดสีด้วยจุดที่อยู่ภายในปริภูมิ 3 มิติโดยจะมีแกนอ้างอิงสำหรับจุดสีนั้นจะอยู่ในปริภูมิซึ่งแต่ละแกนจะมีความเป็นอิสระต่อกันตัวอย่างเช่นในระบบสี RGB จะมีแกนสีคือแกนสีแดง เขียว และน้ำเงิน ในระบบสี HSV จะมีแกนเป็นค่าสี (Hue) ความสว่างสี (Value) และความบริสุทธิ์ของสี (Saturation) ตัวอย่างระบบสีที่นิยมใช้กัน ได้แก่ระบบ RGB, HSL (Hue Saturation Lightness) เป็นต้น [2]

2.2 ระบบสี HSV

ระบบ HSV (Hue Saturation Value) เป็นการพิจารณาโดยใช้ Hue Saturation และ Value คือค่าสีของสีหลัก แดง เขียว และสีน้ำเงิน ในทางปฏิบัติจะอยู่ระหว่าง 0 และ 255 ซึ่งถ้า Hue มีค่าเท่ากับ 0 จะแทนสีแดง และเมื่อ Hue มีค่าเพิ่มขึ้นเรื่อยๆสีก็จะเปลี่ยนแปลงไปตามสเปกตรัมของสีจนถึง 256 จึงจะกลับมาเป็นสีแดงอีกครั้งซึ่งสามารถแทนให้อยู่ในรูปขององศาได้ ดังนี้คือ สีแดงเท่ากับ 0 องศา สีเขียวเท่ากับ 120 องศา สีน้ำเงินเท่ากับ 240 องศา โดยลักษณะโมเดลของระบบ Hue นั้นพบว่าจะมีค่าอย่างน้อยหนึ่งค่าเท่ากับ 0 แต่ถ้ามีสองค่าเท่ากับ 0 แล้วค่า Hue จะเป็นมุมของสี (ค่าสี) มีค่าเป็นไปตามสีที่สามและ ถ้ามีค่าเท่ากับ 0 แล้วจะทำให้ไม่มีค่าของ Hue หรือสีที่ได้จะมีค่าเท่ากับสีขาว ตัวอย่างเช่น จอภาพขาวดำถ้ามีสีใดสีหนึ่งมีค่าเท่ากับ 0 จะทำให้ค่าสีที่ได้เป็นไปตามสีที่เหลือ การให้น้ำหนักในการพิจารณาเมื่อสีแดงมีค่าเท่ากับ 0 ถ้าค่าความบริสุทธิ์ของสีซึ่งถ้ามีค่าเท่ากับ 0 แล้วสีที่ได้จะไม่มี Hue ซึ่งจะเป็นสีขาวล้วนแต่ถ้าค่าความบริสุทธิ์มีค่าเท่ากับ 255 แสดงว่าจะไม่มีแสงสีขาวผสมอยู่เลย ค่าความบริสุทธิ์สามารถคำนวณได้จากความสว่างของสี ซึ่งสามารถวัดได้ด้วยค่าความเข้มของความสว่างแต่ละสีที่ประกอบกัน [2]

2.3 องค์ประกอบทางเคมีของมะเขือเทศ

ในผลมะเขือเทศมีสารจำพวก แคลโรทีนอยด์ ชื่อ ไลโคปีนซึ่งเป็นสารสีแดง และวิตามินหลายชนิด เช่น วิตามินบี 1 บี 2

วิตามินเค โดยเฉพาะวิตามินเอ และวิตามินซี มีในปริมาณสูง มีกรดมาลิก กรดซิตริก ซึ่งให้รสเปรี้ยว และมีกลูตามิก ซึ่งเป็นกรดอะมิโน ช่วยเพิ่มรสชาติให้อาหาร นอกจากนี้ยังประกอบด้วยสารเบต้าแคโรทีน และแร่ธาตุหลายชนิด เช่น แคลเซียม ฟอสฟอรัส เหล็ก เป็นต้น โดยปัจจุบันได้มีการนำมะเขือเทศมาทดลองหาสารเคมี โดยเฉพาะสารเคมีที่สำคัญที่เรียกว่า สารไลโคปีน ในงานวิจัยนี้จะศึกษาและวิเคราะห์ปริมาณสารไลโคปีนเท่านั้น [3,4]

2.3 โครงข่ายประสาทเทียม

เป็นหนึ่งในเทคนิคของการทำเหมืองข้อมูล คือแบบจำลองทางคณิตศาสตร์ สำหรับประมวลผลสารสนเทศด้วยการคำนวณแบบคอนเนกชันนิสต์ เพื่อจำลองการทำงานของเครือข่ายประสาทในสมองมนุษย์ ด้วยวัตถุประสงค์ที่จะสร้างเครื่องมือซึ่งมีความสามารถในการเรียนรู้การจดจำรูปแบบ และการสร้างความรู้ใหม่ เช่นเดียวกับความสามารถที่มีในสมองมนุษย์ แนวคิดเริ่มต้นของเทคนิคนี้ได้มาจากการศึกษาโครงข่ายไฟฟ้าชีวภาพในสมอง ซึ่งประกอบด้วย เซลล์ประสาท หรือนิวรอน (Neurons) และจุดประสานประสาท (Synapses) แต่ละเซลล์ประสาทประกอบด้วยปลายในการรับกระแสประสาทเรียกว่า เดนไดรต์ (Dendrite) ซึ่งเป็นอินพุต และปลายในการส่งกระแสประสาทเรียกว่าแอกซอน (Axon) ซึ่งเป็นเหมือนเอาต์พุตของเซลล์ เซลล์เหล่านี้ทำงานด้วยปฏิกิริยาไฟฟ้าเคมี เมื่อมีการกระตุ้นด้วยสิ่งเร้าภายนอกหรือกระตุ้นด้วยเซลล์ด้วยกัน กระแสประสาทจะวิ่งผ่านเดนไดรต์เข้าสู่นิวเคลียสซึ่งจะเป็นตัวตัดสินใจว่าต้องกระตุ้นเซลล์อื่น ๆ ต่อหรือไม่ ถ้ากระแสประสาทแรงพอ นิวเคลียสก็จะกระตุ้นเซลล์อื่น ๆ ต่อไปผ่านทางแอกซอน สำหรับในคอมพิวเตอร์นิวรอนประกอบด้วยอินพุตและเอาต์พุตเหมือนกัน โดยจำลองให้อินพุตแต่ละอันมีตัวกำหนดน้ำหนักของอินพุต (Weight) โดยนิวรอนแต่ละหน่วยจะมีค่าการกระตุ้นน้อยสุดที่ทำให้เกิดปฏิกิริยาการเคลื่อนไหว (Threshold) เป็นตัวกำหนดว่าน้ำหนักรวมของอินพุตต้องมากขนาดไหนจึงจะสามารถส่งเอาต์พุตไปยังนิวรอนอื่นได้ เมื่อนำนิวรอนแต่ละหน่วยมาต่อกันให้ทำงานร่วมกันการทำงานนี้ในทางตรรกะแล้วก็จะเหมือนกับปฏิกิริยาเคมีที่เกิดในสมอง การทำงานของโครงข่ายประสาทเทียมคือ

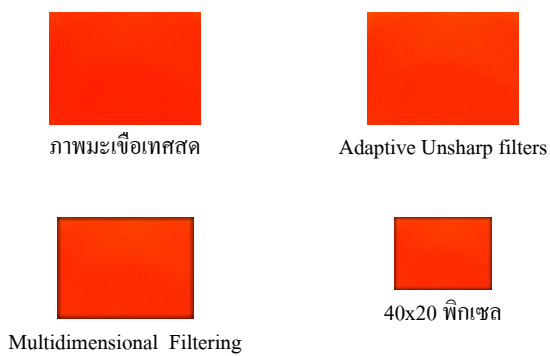
เมื่อมีอินพุตเข้ามายัง โครงข่ายแล้วนำอินพุตมาคูณกับค่าน้ำหนักของแต่ละส่วนผลที่ได้จากอินพุตทุกจุดของนิวรอนจะเอามารวมกันแล้วก็เอามาเทียบกับค่าการกระตุ้นน้อยสุดที่ทำให้เกิดปฏิกิริยาการเคลื่อนไหวที่กำหนดไว้ ถ้าผลรวมมีค่ามากกว่าค่าการกระตุ้นน้อยสุดที่ทำให้เกิดปฏิกิริยาการเคลื่อนไหวแล้ว นิวรอนจะส่งเอาต์พุตออกไปเอาต์พุตนี้ก็จะถูกส่งไปยังอินพุตของนิวรอนอื่นๆที่เชื่อมกันในโครงข่าย ถ้าค่าน้อยกว่าค่าการกระตุ้นน้อยสุดที่ทำให้เกิดปฏิกิริยาการเคลื่อนไหวก็จะไม่เกิดเอาต์พุต สิ่งสำคัญคือต้องทราบค่าน้ำหนักและค่าการกระตุ้นน้อยสุดที่ทำให้เกิดปฏิกิริยาการเคลื่อนไหวสำหรับสิ่งที่เราต้องการเพื่อให้คอมพิวเตอร์รู้จัก ซึ่งเป็นค่าที่ไม่แน่นอน แต่สามารถกำหนดให้คอมพิวเตอร์ปรับค่าเหล่านั้นได้โดยการสอนให้รู้จักรูปแบบ (Pattern) ของสิ่งที่เราต้องการให้รู้จักเรียกว่า Back propagation ซึ่งเป็นกระบวนการย้อนกลับของการรู้จัก ใน Feed-Forward Neural Networks จะมีการใช้อัลกอริทึมแบบ Back-propagation เพื่อใช้ในการปรับปรุงน้ำหนักของเครือข่ายหลังจากใส่รูปแบบข้อมูลสำหรับฝึกให้แก่เครือข่ายในแต่ละครั้งแล้ว ค่าที่ได้รับ จากเครือข่ายจะถูกนำไปเปรียบเทียบกับผลที่คาดหวัง แล้วทำการคำนวณหาค่าความผิดพลาด ซึ่งค่าความผิดพลาดนี้จะถูกส่งกลับเข้าสู่เครือข่ายเพื่อใช้แก้ไขค่าน้ำหนักคะแนนต่อไป [5]

3. ขั้นตอนการทดลองและผลการทดลอง

3.1 หลักการตรวจวิเคราะห์ค่าอัตราส่วนสีจากภาพถ่าย

งานวิจัยนี้ใช้กล้องซีซีดีถ่ายภาพตัวอย่างมะเขือเทศสด เพื่อบันทึกเป็นภาพต้นฉบับ และวิเคราะห์ค่าอัตราส่วนสีในแกนสีแดง สีเขียวและสีน้ำเงิน โดยภาพถ่ายต้นฉบับจะมีขนาดภาพเท่ากับ 640x480 พิกเซล ขนาด 8 บิต ประกอบไปด้วยแกนสีแดง (R) สีเขียว (G) สีน้ำเงิน (B) โดยทำการถ่ายภาพทุกครั้งนั้นจะทำในกล่องทึบแสงและถูกควบคุมค่าความสว่างด้วยหลอดแอลอีดีที่อยู่ในตัวกล้อง โดยคุณสมบัติของหลอดแอลอีดีจะเป็นหลอดไฟฟ้ายุคใหม่ซึ่งจะไม่มีการกระพริบเหมือนหลอดฟลูออโรเรสเซนต์ เพื่อเป็นการรักษาคุณภาพของแสงไม่ให้เกิดข้อผิดพลาด จากนั้นนำภาพถ่ายเข้าสู่กระบวนการประมวลผลทางภาพเพื่อปรับปรุงคุณภาพของภาพถ่ายให้มีค่าคุณภาพสูงขึ้น โดยขั้นตอนแรกเป็นการนำภาพต้นฉบับมาทำการปรับปรุง

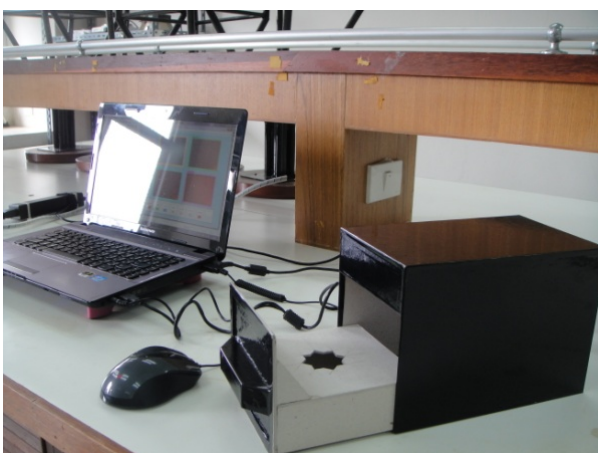
คุณภาพด้วยวิธี อะแดปทีฟอันชาร์ปฟิวส์เตอร์ (Adaptive Unsharp Filters) [5] ซึ่งกระบวนการดังกล่าวจะทำให้จุดสีแต่ละจุดสีที่อยู่ในภาพมีความคมชัดมากยิ่งขึ้น แต่ภาพที่ได้ยังคงมีคุณภาพต่ำเนื่องจากจุดสีของภาพบางจุดสีมีลักษณะเป็นสีขาวปะปนอยู่ อันเนื่องมาจากค่าการสะท้อนแสงที่ไม่เท่ากันซึ่งเกิดจากบริเวณนั้นมีโมเลกุลของน้ำปะปนอยู่ ดังนั้นจึงต้องนำภาพที่ได้มาผ่านกระบวนการฟิลเตอร์อีกครั้งด้วยวิธีมัลติไดเมนชันแนลฟิวเตอร์ (Multidimensional Filtering) [6] ซึ่งทำให้ภาพที่ผ่านกระบวนการดังกล่าวมีคุณภาพสูงขึ้น ขั้นตอนการวิเคราะห์ภาพทั้งหมดแสดงดังภาพที่ 1



ภาพที่ 1: ภาพจากกระบวนการประมวลผลทางภาพ

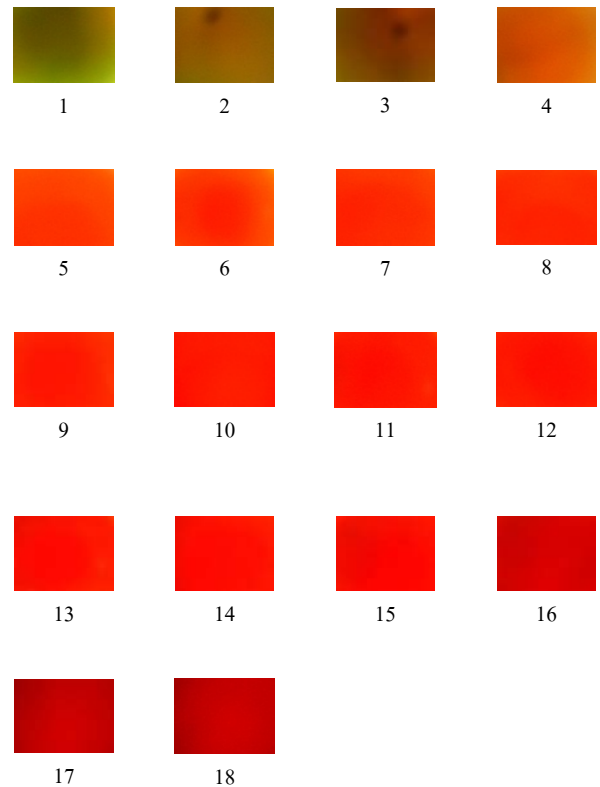
3.2 ขั้นตอนการทดลอง

ในการถ่ายภาพตัวอย่างมะเขือเทศที่ใช้ในการทดลองนี้จะถ่ายภาพในกล่องทึบแสงที่ทำขึ้นมาเฉพาะ ซึ่งในการทดลองครั้งนี้ได้กำหนดจุดอ้างอิงของมะเขือเทศไว้ที่ก้นของมะเขือเทศ



ภาพที่ 2: กล่องทึบแสงสำหรับถ่ายภาพ

โดยเลือกตัวอย่างมะเขือเทศมา 18 ตัวอย่าง โดยแต่ละตัวอย่างจะมีสีที่แตกต่างกันเรียงจากสีแดงอ่อนไปสีแดงเข้ม จากนั้นนำแต่ละตัวอย่างมาถ่ายภาพผ่านกระบวนการข้างต้น ซึ่งจะได้ข้อมูลสี RGB ออกมา แล้วทำการบันทึกผลภาพถ่ายมะเขือเทศที่ได้ ดังภาพที่ 3



ภาพที่ 3: ตัวอย่างมะเขือเทศที่มีสีแตกต่างกันจำนวน 18 ตัวอย่าง

จากนั้นนำมะเขือเทศทั้ง 18 ตัวอย่างมาทำการวัดทางเคมีเพื่อหาปริมาณไลโคปีนด้วยเครื่อง UV-spectrophotometer



ภาพที่ 4: เครื่อง UV-spectrophotometer

จากข้อมูลที่ได้ 18 ข้อมูล ทำการสุ่มเลือกข้อมูลเพื่อใช้สำหรับการสอน (Train) 12 ข้อมูล และชุดข้อมูลที่เหลือใช้สำหรับการทดสอบ (Test) อีก 6 ข้อมูล โดยข้อมูลทั้ง 18 ข้อมูล แสดงในตารางที่ 1

ตารางที่ 1: ข้อมูลค่าสี RGB และปริมาณไลโคปีน

ลำดับ	ค่าสี RGB			Lycopene (mg/100mg)
	R	G	B	
1	138.622	124.593	1.49194	0.439
2	127.248	107.36	1.67689	1.072
3	147.905	69.0104	74.4513	1.263
4	241.385	99.8012	1.66936	1.430
5	244.649	71.811	0.765306	1.542
6	244.649	61.9443	0.798353	1.688
7	244.649	53.5594	0.786969	1.391
8	244.649	47.6932	0.8230.7	1.432
9	244.649	35.452	0.770046	2.514
10	244.649	29.6119	0.413958	2.955
11	244.649	26.0308	0.961117	3.227
12	244.649	24.3432	0.874896	4.134
13	244.649	19.8304	1.04947	4.684
14	244.649	16.5993	0.794619	4.534
15	244.649	11.4994	0.609368	5.222
16	244.244	1.66686	0.337821	7.010
17	230.389	0.519417	0.514215	7.386
18	231.345	0.531393	0.386829	8.184

โครงข่ายประสาทเทียมที่ใช้ในงานวิจัยนี้เลือกใช้ตัวแบบ Multi Layer Feed Forward Back Propagation Neuron Network โดยมีเงื่อนไขที่ใช้ในงานวิจัยดังนี้

- จำนวนอินพุต (Input) มี 3 ค่าคือค่าสี RGB
- จำนวนเอาต์พุต (Output) มี 1 ค่าคือปริมาณไลโคปีน
- จำนวนชั้นซ่อนที่ใช้ทดสอบเริ่มต้นที่ 1 ถึง 6 ชั้นซ่อน
- จำนวน โหนดในแต่ละชั้นซ่อนที่ใช้ทดสอบเริ่มต้นที่ 3 ถึง 18 โหนด
- ชั้นซ่อนใช้ฟังก์ชันถ่ายโอน Tangent Sigmoid ชั้นผลลัพธ์ใช้ฟังก์ชันถ่ายโอนเชิงเส้น
- ฟังก์ชันที่ใช้วัดความเหมาะสมของตัวแบบ (Performance Function) คือค่ากลางของความคลาดเคลื่อนกำลังสอง (Mean Square Error; MSE)
- กำหนดรอบการคำนวณ 1,000 รอบ

3.3 การวัดประสิทธิภาพ

งานวิจัยนี้ได้ทำการวัดประสิทธิภาพของงานวิจัยซึ่งแสดงในรูปแบบของค่าความคลาดเคลื่อน 3 ค่าคือ ค่ากลางของความคลาดเคลื่อนสัมบูรณ์ (Mean Absolute Error; MAE) ค่ากลางของความคลาดเคลื่อนกำลังสอง (Mean Square Error; MSE) และค่ากลางของเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์ (Mean Absolute Percentage Error; MAPE) ตาม (1), (2) และ (3)

$$MAE = \frac{\sum_{i=1}^n |L_i - \hat{L}_i|}{n} \tag{1}$$

$$MSE = \frac{\sum_{i=1}^n (L_i - \hat{L}_i)^2}{n} \tag{2}$$

$$MAPE = \left(\frac{1}{n}\right) \sum_{i=1}^n \left| \frac{L_i - \hat{L}_i}{L_i} \right| \tag{3}$$

เมื่อ L_i คือปริมาณไลโคปีนครั้งที่ i ที่ได้จากการวัด
 \hat{L}_i คือปริมาณไลโคปีนครั้งที่ i ที่ได้จากการพยากรณ์
 n คือจำนวนข้อมูล

จากการทดสอบพบว่าโครงสร้างที่เหมาะสมที่สุดมีโครงข่ายประสาทเทียมแบบ 3-6-1 โดยสามารถแสดงค่าความคลาดเคลื่อนทั้ง 3 ค่า ตามตารางที่ 2

ตารางที่ 2: ค่าความคลาดเคลื่อนจากการทดลอง

MAE	0.6520
MSE	0.3341
MSPE	0.4762

สรุป

จากผลการทดลองวิเคราะห์ค่าความคลาดเคลื่อนในการหาปริมาณสารไลโคปีนในตัวอย่างมะเขือเทศตามตารางที่ 2 พบว่าค่ากลางของความคลาดเคลื่อนสัมบูรณ์ 0.6520 ค่ากลางของความคลาดเคลื่อนกำลังสอง 0.3341 และค่ากลางของเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์ 0.4762 ปัญหาที่พบส่วนหนึ่งที่ทำให้มีค่าความคลาดเคลื่อนเกิดขึ้นเนื่องจากมะเขือเทศมีสีไม่เท่ากันทุกด้านทำให้ตำแหน่งที่กล้องจับภาพได้มีความคลาดเคลื่อน จากผลการวิจัยทั้งหมดสามารถสรุปได้ว่า การใช้กล้องซีซีดีในการตรวจวัดปริมาณไลโคปีนในมะเขือเทศด้วยโครงข่ายประสาทเทียมสามารถตรวจวัดปริมาณไลโคปีนในมะเขือเทศได้อย่างถูกต้อง มีความสะดวก รวดเร็ว และราคาประหยัดอีกทั้งไม่ต้องทำลายมะเขือเทศที่นำมาใช้ทดสอบและสามารถแก้ปัญหาในเรื่องการวัดทางเคมีที่ต้องใช้ผู้เชี่ยวชาญในการวิเคราะห์ซึ่งเป็นไปตามเป้าหมายของการวิจัยครั้งนี้

กิตติกรรมประกาศ

งานวิจัยนี้ได้รับทุนอุดหนุนวิจัยจากคณะวิทยาศาสตร์และวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตเฉลิมพระเกียรติ จังหวัดสกลนคร

เอกสารอ้างอิง

[1] รสสุคนธ์ พุ่มพันธุ์วงศ์ 2548, "พืชผักสวนครัวอินทรีย์", มะเขือเทศ. หน้า 125- 129.
 [2] . <http://fivedots.coe.psu.ac.th/~montri/Teaching/image/chap1.htm>
 [3] ชัยสิทธิ์ ชินวัตร และคณะ .“การปลูกมะเขือเทศปลอดสารพิษและวิธี

เพิ่มผลผลิตลักษณะทางพฤกษศาสตร์คุณภาพทางโภชนาการ ,”พันธุ์มะเขือเทศ และหลักการเลือกมะเขือเทศ หน้า.9-14, 19-24.
 [4] ปริญญา สงวนศักดิ์ คู่มือ MATLAB ฉบับสมบูรณ์. บริษัทไอดีซี พรีเมียร์ จำกัด 2533.
 [5] <http://www.gotoknow.org/posts/163433>
 [6] A.Polesel,G. Ramponi,V.J.Mathews, “Adaptive unsharp masking for contrast enhancement”, IEEE Trans. On Image Processing Vol. 1, No.3, 26-29 Oct 1997, pp.267-270.
 [7] Tsuhan Chen, Member, P.P. Vaidyanathan, “Vector Space Framework For Unification of One-and Multidimensional Filter Bank The”,IEEE Trans. On sine Processing, Vol. 42, No. 8 Aug 1994, pp. 2006-2021.
 [8] Minsky and Papert Perceptrons, An introduction to computational geometry, MIT press, expanded edition. 1969.
 [9] Aleksander, I. and Morton, H. An introduction to neural computing. 2nd edition.
 [10]http://www.thapra.lib.su.ac.th/objects/thesis/fulltext/snamcn/Jaravee_Chantasitiporn/Fulltext.pdf

การเพิ่มประสิทธิภาพดัชนีบิตแมปแบบคู่กันโดยใช้วิธีการเข้ารหัส

Enhancing Dual Bitmap Index with Efficient Encoding

ศินเตร กิมเส็ง(Sinate Kimseng)¹ และ ศิริรัตน์ วมัชโยบล(Sirirut Vanichayobon)²

ห้องปฏิบัติการวิจัยเทคโนโลยีระบบสารสนเทศและการประยุกต์ (iSTAR Lab) ภาควิชาวิทยาการคอมพิวเตอร์

คณะวิทยาศาสตร์ มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตหาดใหญ่

¹5510220103@email.psu.ac.th, ²sirirut.v@psu.ac.th

บทคัดย่อ

คลังข้อมูลเป็นที่เก็บข้อมูลสำหรับสนับสนุนการตัดสินใจของผู้บริหาร ลักษณะการสอบถามที่เกิดขึ้นบ่อย ๆ ในคลังข้อมูลเป็นการสอบถามแบบทันทีทันใด ดัชนีบิตแมปเป็นเทคนิคที่นิยมใช้ค้นหาข้อมูลที่ต้องการ เนื่องจากการดำเนินการ Boolean ที่มีประสิทธิภาพของ AND OR และ NOT อย่างไรก็ตามดัชนีบิตแมปเหมาะสมสำหรับแอททริบิวต์ที่มีคาร์ดินอลิตีต่ำเท่านั้น บทความนี้เสนอเทคนิคการเพิ่มประสิทธิภาพดัชนีบิตแมปแบบคู่กัน โดยใช้วิธีการเข้ารหัสที่เหมาะสมทำให้สามารถนำไปใช้กับแอททริบิวต์ที่มีคาร์ดินอลิตีสูงซึ่งขณะที่ยังคงรักษาประสิทธิภาพในการสอบถาม

คำสำคัญ: คลังข้อมูล ดัชนีบิตแมป ดัชนีบิตแมปแบบคู่กัน

Abstract

A data warehouse is a repository data for supporting decision makers. Most of queries against the data warehouse are ad hoc query. Bitmap index is a popular technique used to retrieve target records because of the efficient Boolean operation AND, OR and NOT on bits. However, bitmap index is suitable for only low cardinality attributes. This paper proposes a technique to enhance the Dual bitmap index to optimize the use of space in high cardinality attributes while maintaining query processing time performance.

Keyword: Data Warehouse, Bitmap Index, Dual Bitmap Index

1. บทนำ

คลังข้อมูล (Data Warehouse) เป็นที่ใช้เก็บรวบรวมข้อมูลให้อยู่ในที่เดียวกัน มีการจัดโครงสร้างตามเนื้อหาที่เราสนใจ [1] โดยข้อมูลที่เก็บรวบรวมนั้นจะเป็นข้อมูลตั้งแต่อดีตจนถึงปัจจุบันเพื่อสนับสนุนการตัดสินใจ [2] ในระบบคลังข้อมูลการประมวลผลข้อมูลจะเป็นแบบ OLAP (Online Analytical Processing)[3] และแบบทันทีทันใด (Ad Hoc) และการสอบถามส่วนใหญ่ซับซ้อนทำให้การประมวลผลข้อมูลต้องใช้เวลาาน จากปัญหาที่กล่าวมาจึงมีการทำงานวิจัยที่เกี่ยวข้องกับคลังข้อมูลมากมายส่วนใหญ่แล้วจะเน้นสองเรื่องหลัก คือ เรื่องของพื้นที่ที่ใช้และเรื่องของเวลาในการเข้าถึงข้อมูล เมื่อก้าวถึงการเข้าถึงข้อมูล วิธีการเพิ่มความเร็วในการเข้าถึงข้อมูลมีหลายวิธีด้วยกันยกตัวอย่างเช่น วิธีการ Hashing เหมาะกับการเข้าถึงโดยตรง การประมวลผลแบบ Parallel เป็นการเพิ่ม Hardware เพื่อให้ทำงานเร็วขึ้นและ วิธีการทำดัชนี เป็นวิธีการเพิ่มความเร็วในการเข้าถึงข้อมูลโดยไม่ต้องเสียค่าใช้จ่ายใด ๆ แต่เป็นการใช้พื้นที่แลกกับเวลาในการสอบถามที่เร็วขึ้น

การทำดัชนีมีหลายแบบ เช่น B-Tree, B+ Tree และดัชนีบิตแมป ในบทความนี้จะกล่าวถึงเฉพาะดัชนีบิตแมปเนื่องจากดัชนีบิตแมปมีความนิยมในการใช้ในการประมวลผลแบบทันทีทันใด [1, 3, 4] เพราะสามารถดำเนินการระดับบิต (AND, OR, XOR, NOT) ก่อนจะเข้าถึงข้อมูลจริง

การทำดัชนีบิตแมป (Bitmap Index) ในงานวิจัยที่ผ่านมา มีการทำดัชนีบิตแมปหลายชนิดโดยแบ่งเป็น 2 กลุ่ม คือ กลุ่มที่ใช้วิธีการบีบอัด และกลุ่มที่ใช้วิธีการลงรหัส ยกตัวอย่างดัชนีบิตแมปที่ใช้วิธีการลงรหัส 5 ชนิด คือ ดัชนีบิตแมปแบบพื้นฐาน (Simple Bitmap Index) [4] ดัชนีบิตแมปแบบช่วง (Interval Bitmap Index) [5] ดัชนีบิตแมปแบบกระจาย (Scatter Bitmap Index) [6] ดัชนีบิตแมปแบบเข้ารหัส (Encoded Bitmap Index) [7, 8] และ ดัชนีบิตแมปแบบคู่กัน (Dual Bitmap Index) [9]

จากการศึกษาในงานวิจัยที่ผ่านมาผู้วิจัยสนใจการดัชนีบิตแมปแบบคู่กัน เพราะยังมีข้อบกพร่องที่สามารถแก้ไขได้ในเรื่องของพื้นที่จัดเก็บดัชนี ผู้วิจัยจึงคิดค้นวิธีการเพิ่มประสิทธิภาพให้กับดัชนีบิตแมปแบบคู่กันขึ้น โดยเล็งเห็นข้อดีและข้อเสียของดัชนีบิตแมปแบบคู่กันคือ หนึ่งแอททริบิวต์จะแทนด้วย 2 บิตแมปเวกเตอร์ ส่วนข้อเสียของดัชนีบิตแมปแบบคู่กันคือขนาดพื้นที่จัดเก็บดัชนีบิตแมปยังใช้พื้นที่มาก ดังนั้นจึงนำเสนอวิธีการเพิ่มประสิทธิภาพดัชนีบิตแมปแบบคู่กันเพื่อลดพื้นที่ในการจัดเก็บดัชนีให้น้อยลงโดยใช้วิธีการเข้ารหัส และใช้เวลาการสอบถามที่เท่ากันในบางกรณี

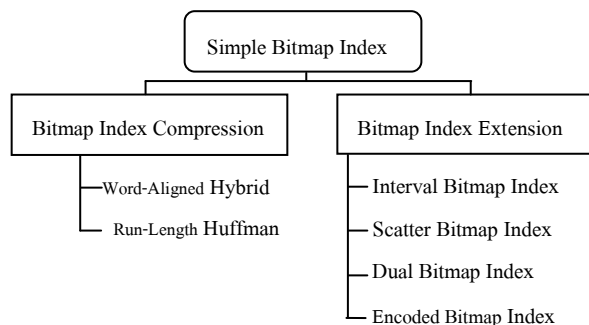
เอกสารนี้ประกอบด้วย 5 ส่วน ส่วนที่ 2 กล่าวถึงดัชนีบิตแมปแบบคู่กัน ส่วนที่ 3 วิธีการที่นำเสนอ ส่วนที่ 4 ผลการวิจัยและ ส่วนที่ 5 ส่วนสรุปและวิจารณ์ผล

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ดัชนีบิตแมป

ดัชนีบิตแมปเป็นเทคนิคที่ช่วยให้เราสามารถค้นหาข้อมูลให้เร็วขึ้น โดยตารางดัชนีจะแทนด้วยบิต 0 และ 1 ดัชนีบิตแมปแบบพื้นฐานมีการแทนค่าข้อมูล 1 บิตแมปเวกเตอร์ต่อ 1 ค่าของแอททริบิวต์ซึ่ง S^i มีค่าเท่ากับ 1 เมื่อเรคอร์ด $i = v$ ข้อดีคือใช้เวลาสอบถามน้อย ข้อเสียคือ ถ้าคาร์ดินอลิตี้ C มากจะใช้พื้นที่มากดังนั้นจึงได้มีการพัฒนาเทคนิคเพิ่มประสิทธิภาพของดัชนีบิตแมปแบบพื้นฐานแบ่งเป็น 2 กลุ่มคือ 1. การบีบอัด (Bitmap Index Compression) และ 2. การลงรหัส (Bitmap Index Extension) ดังแสดงในภาพที่ 1

กลุ่มที่ใช้วิธีการบีบอัด ประกอบด้วย WAH (Word – Aligned Hybrid) [10] และ RLH (Run – Length Huffman) [11] ข้อดีก็คือ ลดพื้นที่ที่เก็บดัชนี แต่ข้อเสียก็คือ เวลาสอบถามจะต้องแปลงกลับไปให้อยู่ในรูปแบบของดัชนีบิตแมปแบบพื้นฐานและสอบถามแบบเดียวกับดัชนีบิตแมปแบบพื้นฐาน



ภาพที่ 1: การเพิ่มประสิทธิภาพในเชิงพื้นที่ของดัชนีบิตแมป [8]

กลุ่มที่ใช้วิธีการลงรหัส ประกอบด้วย 1) ดัชนีบิตแมปแบบช่วง [5] เป็นดัชนีบิตแมปที่ใช้พื้นที่ได้มีประสิทธิภาพกว่าดัชนีบิตแมปแบบพื้นฐาน การสอบถามข้อมูล 2 บิตแมปและมีโอกาสตรวจสอบ 1 บิตแมปเวกเตอร์เมื่อคาร์ดินอลิตี้น้อยกว่าหรือเท่ากับ 3 2) ดัชนีบิตแมปแบบกระจาย[6] ใช้พื้นที่ในการจัดเก็บดัชนีน้อยกว่าดัชนีบิตแมปแบบช่วง การสอบถามข้อมูลจะดำเนินการกับ 2 บิตแมปเวกเตอร์ 3) ดัชนีบิตแมปแบบคู่กัน [9] ใช้บิตแมปเวกเตอร์น้อยกว่าดัชนีบิตแมปแบบกระจาย การสอบถามข้อมูลจะดำเนินการระหว่าง 2 บิตแมปเวกเตอร์ และ 4) ดัชนีบิตแมปแบบเข้ารหัส [7, 8] เป็นดัชนีบิตแมปที่ใช้พื้นที่น้อยที่สุด แต่การสอบถามข้อมูลแบบดำเนินการกับทุกบิตแมปเวกเตอร์ ข้อดีของดัชนีบิตแมปกลุ่มที่ 2 ใช้วิธีการลงรหัสคือการสอบถาม (Query) สามารถทำได้ทันที และมีประสิทธิภาพในเรื่องของการลดพื้นที่การจัดเก็บดัชนีพอสมควร

2.2 ดัชนีบิตแมปแบบคู่กัน (Dual Bitmap Index)

ดัชนีบิตแมปแบบคู่กันเป็นเทคนิคที่นำเสนอโดย Wattanakitrungrroj และ Vanichayobon [9] วิธีการสร้างดัชนีบิตแมปแบบคู่กันมีขั้นตอนดังนี้ (กำหนดให้ C แทนค่าคาร์ดินอลิตี้)

1. กำหนดหาจำนวนบิตที่ใช้แทนค่าแต่ละค่า

$$n = \left\lceil \sqrt{2C + 0.25} + 0.5 \right\rceil \quad (1)$$

2. กำหนดค่า 0 ถึง C-1 ให้แก่ค่าของแอททริบิวต์
3. กำหนดตำแหน่งบิตที่ต้องให้ค่าเท่ากับ 1 โดยใช้สมการ

$$r_i = \left\lceil \sqrt{2v_i + 2.25} - 0.5 \right\rceil \quad (2)$$

$$s_i = v_i - \frac{r(r-1)}{2} \quad (3)$$

ตัวอย่างการสร้างดัชนีบิตแมปแบบกลุ่มบนแอททริบิวต์ X ที่คาร์ดินอลลิตี้ C=16 (คือค่า A ถึง P) จะได้ n = 7 กำหนดค่าตัวเลขให้กับ A ถึง P เช่น A=0, B=1 เป็นต้น ตารางที่ 1 แสดงค่าของแอททริบิวต์ X

ตารางที่ 1: ตารางแสดงค่าของแอททริบิวต์ X

X	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
v	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

ถ้าต้องการลรหค่า F (v=5) จากสมการที่ (2) และ (3) ได้ r = 3 และ s = 2 ดังนั้นสองบิตที่เท่ากับ 1 คือตำแหน่งที่ 3 (D³) และ ตำแหน่งที่ 2 (D²) ดังที่แสดงในตารางที่ 2(ข)

ตารางที่ 2: ตัวอย่างการลรหค่าดัชนีบิตแมปแบบกลุ่ม (C=16)

RID	X	D ⁶	D ⁵	D ⁴	D ³	D ²	D ¹	D ⁰	D ³ ∧ D ²
1	K	0	1	0	0	0	0	1	0
2	F	0	0	0	1	1	0	0	1
3	D	0	0	0	1	0	0	1	0
4	P	1	0	0	0	0	0	1	0
5	N	0	1	0	1	0	0	0	0
6	B	0	0	0	0	1	0	1	0
7	L	0	1	0	0	0	1	0	0
8	E	0	0	0	1	0	1	0	0
9	A	0	0	0	0	0	1	1	0
10	F	0	0	0	1	1	0	0	1

(ก) แอททริบิวต์ (ข) ดัชนีบิตแมปแบบกลุ่ม (ค) ผลลัพธ์

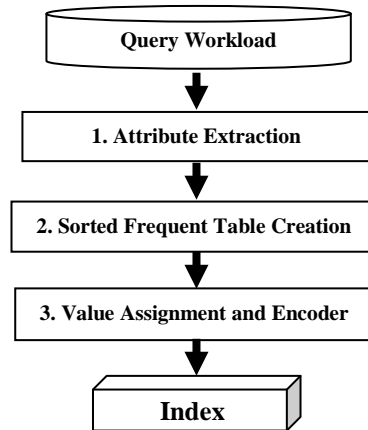
การสอบถามแบบค่าเท่ากันบนดัชนีบิตแมปแบบกลุ่มทำได้โดย D^r ∧ D^s เช่น ถ้าต้องการสอบถามค่า F (v = 5) จากสมการ (2) และ (3) ได้ r = 3 และ s = 2 แสดงว่าจะต้องทำการ AND ระหว่างบิต D³ และ D² ได้ผลลัพธ์ดังภาพที่ 2(ค)

ข้อดีของดัชนีบิตแมปแบบกลุ่มมีการดำเนินการกับบิตแมปเวกเตอร์ 2 บิตเท่านั้นส่วน ข้อเสียคือยังคงใช้พื้นที่ในการจัดเก็บดัชนีมาก

3. Enhancing Dual Bitmap Index

3.1 การสร้าง Enhancing Dual Bitmap Index (EDBI)

การสร้าง EDBI แบ่งเป็น 3 ขั้นตอนคือ 1) Attribute Extraction 2) Sorted Frequent Table Creation และ 3) Value Assignment and Encoder ดังภาพที่ 2



ภาพที่ 2: ขั้นตอนการเพิ่มประสิทธิภาพดัชนีบิตแมปแบบกลุ่ม

1). Attribute Extraction

ค่าแอททริบิวต์จาก Query Workload จะถูกสกัดออกมาในขั้นตอนนี้ ภาพที่ 3 แสดงตัวอย่างของ Query Workload 4 รายการและตารางที่ 3 แสดงผลลัพธ์ที่ได้จากการสกัดค่าในแอททริบิวต์ X จาก Query Workload

Q1: SELECT * FROM T WHERE X IN (A, E, G, O, P, C, M, N)
Q2: SELECT * FROM T WHERE X = A OR X = E
Q3: SELECT * FROM T WHERE X = E OR X = D
Q4: SELECT * FROM T WHERE X = D

ภาพที่ 3: Query Workload

ตารางที่ 3: ตารางค่าของแอททริบิวต์

Query	ค่าของแอททริบิวต์ X
Q1	A, E, G, O, P, C, M, N
Q2	A, E
Q3	E, D
Q4	D

2). Sorted Frequent Table Creation

ในขั้นตอนนี้ค่าความถี่(f) ของค่าของแอททริบิวต์ที่ถูกสกัด แต่ละตัวจะถูกคำนวณและจัดเก็บไว้ในตารางแสดงความถี่ (Frequent Table) จากนั้นเรียงค่าของแอททริบิวต์ X ตามความถี่ จากค่ามากไปหาค่าน้อยและจัดเก็บในตารางจัดเรียงความถี่ (Sorted Frequent Table) ตารางที่ 4 และ 5 แสดง Frequent Table และ Sorted Frequent Table ของแอททริบิวต์ X ตามลำดับ

ตารางที่ 4: Frequent Table ของแอททริบิวต์ X

X	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
f	2	0	1	2	3	0	1	0	0	0	0	0	1	1	1	1

ตารางที่ 5: Sorted Frequent Table ของแอททริบิวต์ X

X	E	A	D	C	G	M	N	O	P	B	H	I	J	L	F	K
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

3). Value Assignment and Encoder

- คำนวณหาจำนวนบิตที่ใช้แทนค่าแต่ละค่า

$$nB = 2 \left\lceil \left\lceil \log_2 \left[\sqrt{2C + 0.25} + 0.5 \right] \right\rceil \right\rceil \quad (4)$$
- สร้างตาราง EDBI Mapping เพื่อลดเวลาในการคำนวณในกระบวนการสร้างดัชนีและการสอบถาม
 - คำนวณหาค่าที่ใช้แทนค่าที่ถูกสอบถามบ่อยสุด

$$V_0 = \frac{((2^{nB/2} - 1) + 0.5)^2}{2} - 1.125 \quad (5)$$
 - กำหนดค่าให้กับค่าที่ถูกสอบถามบ่อยตามลำดับโดยที่

$$V_{i+1} = V_i - 1 \quad \text{for } i = 0, 1, 2, \dots, C - 2 \quad (6)$$
 - นำค่า V_i ที่ได้มาคำนวณหาค่า r^i และ s^i

for $i = 0, 1, \dots, \frac{nB}{2} - 1$ โดยใช้สมการ [9]

$$r^i = \left\lceil \sqrt{2V_i + 2.25} - 0.5 \right\rceil \quad (7)$$

$$s^i = V_i - \frac{r^i(r^i - 1)}{2} \quad (8)$$
- ทำการลงรหัสโดยแทนค่า r^i และ s^i ด้วยเลขฐานสองโดยใช้จำนวนบิตเท่า ๆ กันคือ $\frac{nB}{2}$ บิต

ภาพที่ 4 ขั้นตอนวิธีการกำหนดค่าและการลงรหัส EDBI

ภาพที่ 4 แสดงขั้นตอนวิธีการกำหนดค่า (Value Assignment) และการลงรหัส (Encoder) โดยจำนวนบิต(nB) ที่ใช้แทนค่าแต่ละค่าของแอททริบิวต์จะถูกคำนวณ จากนั้นตาราง EDBI Mapping ซึ่งเก็บค่ากำหนด(V_i) ซึ่งคำนวณจากสมการที่ (5) และ (6) และค่าที่ใช้ลงรหัสบิตแมป (r^i และ s^i) ซึ่งคำนวณสมการที่ (7) และ (8) จะถูกสร้างขึ้น จากนั้นทำการลงรหัสโดยแทนค่า r^i และ s^i ด้วยเลขฐานสองโดยใช้จำนวนบิตที่เท่ากันคือ $\frac{nB}{2}$ บิต

จากตัวอย่างข้างต้นแอททริบิวต์ X ที่มีคาร์ดินอลิตี้ $C=16$ จากสมการ (4) และ(5) จะได้ว่า $nB = 6$ และ $V_0=27$ ตารางที่ 6 แสดงการกำหนดค่าให้กับแอททริบิวต์ X ทั้ง 16 ค่า และค่า r^i และ s^i ที่ได้จากสมการ (7) และ(8)

ตารางที่ 6: EDBI Mapping Table ของแอททริบิวต์ X

X	E	A	D	G	M	N	O	B	C	H	I	J	L	P	F	K
V_i	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12
r^i	7	7	7	7	7	7	7	6	6	6	6	6	6	5	5	5
s^i	6	5	4	3	2	1	0	5	4	3	2	1	0	4	3	2

จากนั้นให้ใช้ 3 บิตในการลงรหัส r และ s เช่นค่า F จากตาราง EDBI Mapping จะได้ $r = 5$ และ $s = 3$ สามารถลงรหัส $r=101$ และ $s=011$ ตารางที่ 7(ข) แสดงการลงรหัส EDBI ของแอททริบิวต์ X จะเห็นว่า EDBI ใช้ 6 บิตแมปเวกเตอร์ซึ่งน้อยกว่าดัชนีบิตแมปแบบคู่กัน (7 บิตแมปเวกเตอร์)

ตารางที่ 7: แสดงการลงรหัสแบบ EDBI

RID	X	R			S		
		r^2	r^1	r^0	s^2	s^1	s^0
1	K	1	0	1	0	1	0
2	F	1	0	1	0	1	1
3	D	1	1	1	1	0	0
4	P	1	0	1	1	0	0
5	N	1	1	1	0	0	1
6	B	1	1	0	1	0	1
7	L	1	1	0	0	0	0
8	E	1	1	1	1	1	0
9	A	1	1	1	1	0	1
10	F	1	0	1	0	1	1

(ก)แอททริบิวต์

(ข) EDBI

3.2 การสอบถามโดยใช้ EDBI

การสอบถามแบบค่าเท่ากันบน EDBI มีขั้นตอนการสอบถามดังแสดงในภาพที่ 5

1. ทำการlookup ค่า r และ s ของค่าที่ต้องการสอบถาม (V) จากตาราง EDBI Mapping
2. กำหนดให้ PATTERN คือ เลขฐานสองของค่า s ให้เรียงต่อจากค่า r
3. $if(s = 0)$
 ทำการจับคู่เรคอร์ดที่ตรงเงื่อนไข
 $(U_s^0 \cup U_s^1 \cup U_s^2 \cup \dots \cup U_s^{(nB/2)-1} = 0)$
 else
 ทำการจับคู่เรคอร์ดที่ตรงเงื่อนไข
 $(\cap r^i = 1) \text{ และ } (\cap s^j = 1) \text{ โดยที่ } (0 \leq i, j \leq \frac{nB}{2} - 1)$
 สำหรับค่าตำแหน่ง i, j ที่มีค่า 1 ใน PATTERN
4. นำ PATTERN มาเทียบกับผลลัพธ์ที่ได้จากข้อ 3

ภาพที่ 5: ขั้นตอนวิธีการสอบถามบน EDBI

ตัวอย่างการสอบถามแบบค่าเท่ากันว่าบนแอทริบิวต์ X มีเรคอร์ดใดบ้างที่มีค่าเท่ากับ F ทำได้ด้วยวิธีการดังนี้

1. ค้นหา r และ s ของ F จากตาราง EDBI Mapping (ตารางที่ 6) ได้ $r = 5$ และ $s = 3$ (ดูตารางที่ 8(ข))
2. สร้าง PATTERN ได้เป็น 101011 ($r = 101$ $s = 011$)
3. $s = 3$ ดังนั้น ทำการจับคู่เรคอร์ดที่ตรงเงื่อนไข
 $((r^2 \cap r^0) = 1) \text{ และ } ((s^1 \cap s^0) = 1)$
4. นำ PATTERN ไปเทียบกับตารางผลลัพธ์ที่ได้จากข้อ 3 คำตอบที่ได้คือแถวที่ 2 และ 10 แสดงดังตารางที่ 8(ก) และ 8(ข)

ตารางที่ 8: แสดงผลลัพธ์จากการสอบถามโดยใช้ EDBI

RID	X	R			S		
		r^2	r^1	r^0	s^2	s^1	s^0
1	K	1	0	1	0	1	0
2	F	1	0	1	0	1	1
3	D	1	1	1	1	0	0
4	P	1	0	1	1	0	0
5	N	1	1	1	0	0	1
6	B	1	1	0	1	0	1
7	L	1	1	0	0	0	0
8	E	1	1	1	1	1	0
9	A	1	1	1	1	0	1
10	F	1	0	1	0	1	1

(ก)แอทริบิวต์

(ข) EDBI

4. ผลการดำเนินการวิจัย

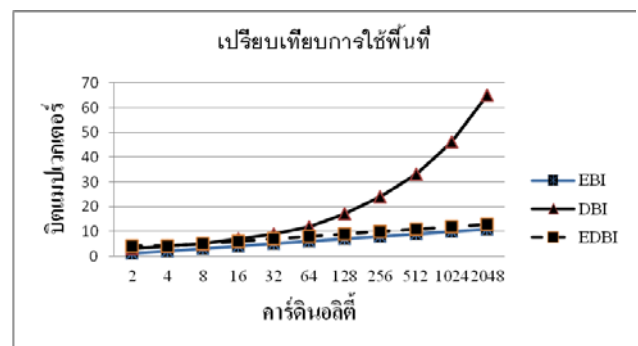
ในส่วนนี้เป็นการวิเคราะห์ประสิทธิภาพในเรื่องของการใช้พื้นที่จัดเก็บดัชนีและประสิทธิภาพในเรื่องของเวลาที่ใช้ในการสอบถาม

4.1 วิเคราะห์ประสิทธิภาพในเรื่องของพื้นที่จัดเก็บดัชนี

ตารางที่ 9 : แสดงการใช้พื้นที่ของดัชนีบีตแมป

ดัชนีบีตแมป	ตัวย่อ	พื้นที่
		จำนวนบีตแมปเวกเตอร์ที่ใช้ในการสร้างดัชนี
Simple Bitmap Index[3]	SBI	C
Interval Bitmap Index[9]	IBI	$\lceil C/2 \rceil$
Scatter Bitmap Index[4]	ScaBI	$\lceil 2\sqrt{C} \rceil$
Dual Bitmap Index[6]	DBI	$\lceil \sqrt{2C + 0.25} + 0.5 \rceil$
Enhancing Dual Bitmap Index	EDBI	$2 \left\lceil \log_2 \left[\sqrt{2C + 0.25} + 0.5 \right] \right\rceil$
Encoded Bitmap Index[5]	EBI	$\lceil \log_2 C \rceil$

ตารางที่ 9 แสดงการใช้พื้นที่ของดัชนีบีตแมป 6 ชนิด พบว่า SBI ใช้พื้นที่จัดเก็บดัชนีมากที่สุด รองลงมาคือ IBI, ScaBI, DBI, EDBI และ EBI ตามลำดับ กราฟในภาพที่ 6 แสดงให้เห็นถึงประสิทธิภาพการใช้พื้นที่ของ EDBI ใกล้เคียง EBI และเหนือกว่า DBI อย่างมากสำหรับแอทริบิวต์ที่มีคาร์ดินอลิตี้สูง

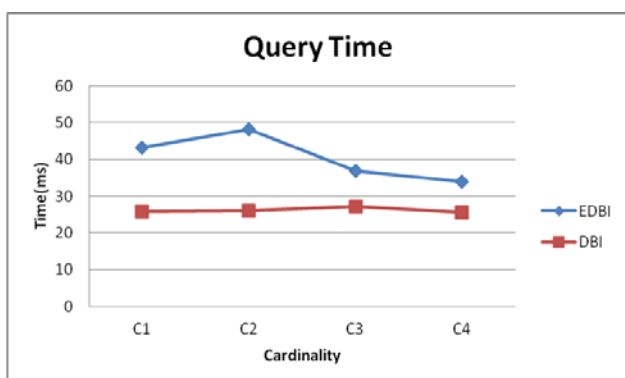


ภาพที่ 6: เปรียบเทียบการใช้พื้นที่ของดัชนีบีตแมป 3 แบบ

4.2 วิเคราะห์ประสิทธิภาพในเรื่องของเวลา

การวิเคราะห์ประสิทธิภาพในเรื่องของเวลาที่ใช้ในการสอบถามใช้ชุดข้อมูลตาราง order เลือกสครมที่มีคาร์ดินอลิตี้ C ต่าง ๆ กันดังนี้ C1=32, C2=64, C3=128 และ C4=256 จาก TPC-H [12] และทำการทดลองเปรียบเทียบเวลา 2 ดัชนีคือ DBI และ EDBI ภาพที่ 7 แสดงการเปรียบเทียบการใช้เวลาในการสอบถามของดัชนีบีตแมป DBI และ EDBI จะเห็นว่า DBI ใช้เวลาน้อยกว่า EDBI ในทุกการสอบถามทั้งนี้เป็นเพราะว่า ในแต่ละการสอบถาม DBI อ่านแค่ 2 บีตแมปเวกเตอร์ ในขณะที่ EDBI อ่านมากกว่า 2 บีตแมปเวกเตอร์ จำนวนบีตแมปเวกเตอร์ที่ EDBI อ่านจะอยู่ระหว่าง 2- nB บีตแมปเวกเตอร์ ส่วน EBI จะอ่าน $\lceil \log_2 C \rceil$ บีตแมปเวกเตอร์

ถ้าพิจารณาประสิทธิภาพในแง่ของพื้นที่และเวลาพร้อมกัน จะเห็นว่า EDBI มีประสิทธิภาพเหนือกว่าดัชนีบีตแมปอื่น ๆ



ภาพที่ 7: เปรียบเทียบการใช้เวลาของดัชนีบีตแมป 2 แบบ

5. สรุป

EDBI เป็นการนำดัชนีบีตแมปแบบคู่กันมาลงรหัสแบบใหม่และใช้ ข้อมูลค่าแอททริบิวต์ที่ถูกสอบถามบ่อย ๆ มาช่วยในการลงรหัสทำให้มีประสิทธิภาพกว่าการลงรหัสแบบดั้งเดิมนำไปสู่การเพิ่มขีดความสามารถของดัชนีแบบคู่กันในการใช้กับแอททริบิวต์ที่มีคาร์ดินอลิตี้สูงในแง่ของพื้นที่ที่ตีเทียบเท่าและในแง่ของเวลาที่ใช้ในการสอบถามโดยเฉลี่ยน้อยกว่าดัชนีแบบเข้ารหัส

สำหรับงานที่จะทำในอนาคตได้แก่ การนำ EDBI ไปใช้กับการลงรหัสเป็นกลุ่มสำหรับการสอบถามแบบสมาชิกเพราะ

EDBI สามารถลดจำนวนการอ่านบีตแมปเวกเตอร์ได้มากกว่าดัชนีบีตแมปแบบคู่กันซึ่งต้องอ่านทีละ 2 บีตแมปเวกเตอร์สำหรับแต่ละค่าในการสอบถามแบบสมาชิก

เอกสารอ้างอิง

- [1] William H. Inmon, *Building the data warehouse.*, Wiley Computer Publishing: New York, 1996.
- [2] Vincent Rainardi, *Building a Data Warehouse: With Examples in SQL Server*, Springer-Verlag: New York, 2008.
- [3] S. Chaudhuri, and U. Dayal. "An overview of data warehousing and OLAP technology.", *ACM Sigmod record* 26.1, pp. 65-74, 1997.
- [4] C.Y. Chan and Y. E. Ioannidis. "Bitmap index design and evaluation.", *ACM SIGMOD Record* 27.2, pp.355-366, 1998.
- [5] C.Y. Chan and Y. E. Ioannidis. "An efficient bitmap encoding scheme for selection queries.", *ACM SIGMOD Record* 28.2, pp. 215-226, 1999.
- [6] S.Vanichayobon, J.Manfuekphan, and L.Gruenwald. "Scatter Bitmap: Space-Time Efficient Bitmap Indexing for Equality and Membership Queries.", *Cybernetics and Intelligent Systems, 2006 IEEE Conference on*. IEEE, 2006.
- [7] M.C. Wu and A.P. Buchmann. "Encoded bitmap indexing for data warehouses.", *Data Engineering, 1998. Proceedings., 14th International Conference on*. IEEE, 1998.
- [8] A. Keawpibal, N. Wattanakitrunroj and S. Vanichayobon. "Enhanced Encoded Bitmap Index for equality query.", *Computing Technology and Information Management (ICCM), 2012 8th International Conference on*. Vol. 1. IEEE, 2012.
- [9] N. Wattanakitrunroj and S. Vanichayobon. "Dual Bitmap Index: Space-Time Efficient Bitmap Index for Equality and Membership Queries.", *Communications and Information Technologies, 2006. ISCIT'06. International Symposium on*. IEEE, 2006.
- [10] F. Delière and T. B. Pedersen. "Position list word aligned hybrid: optimizing space and performance for compressed bitmaps.", *Proceedings of the 13th International Conference on Extending Database Technology*. ACM, 2010.
- [11] M. Stabno and R. Wrembel. "RLH: Bitmap compression technique based on run-length and Huffman encoding.", *Information Systems* 34.4, pp.400-414, 2009.
- [12] Transaction Processing Performance Council (TPC), "TPC-H: An Ad Hoc Decision Support Benchmark", Version 2.14.1, [Online]. Available from: <http://www.tpc.org/tpch.>, [2013, December 11].

การคัดเลือกดัชนี Bitmap Join Index สำหรับคลังข้อมูลโดยมีพื้นที่จัดเก็บดัชนีที่จำกัด

Bitmap Join Index Selection for Data Warehouse under the Space Constraint

หทัย แก้วกรณ์ (Hathai Kaewkorn)¹ และศิริรัตน์ วนิชโยบล (Sirirut Vanichayobon)²

ห้องปฏิบัติการวิจัยเทคโนโลยีระบบสารสนเทศและการประยุกต์ (iSTAR Lab) ภาควิชาวิทยาการคอมพิวเตอร์

คณะวิทยาศาสตร์ มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตหาดใหญ่

¹5510220127@email.psu.ac.th, ²sirirut.v@psu.ac.th

บทคัดย่อ

บทความนี้นำเสนอระบบการคัดเลือกดัชนี Bitmap Join Index ของคลังข้อมูลที่มีโครงสร้างแบบ Star Schema เพื่อเป็นการเพิ่มประสิทธิภาพด้านเวลาในการสอบถามข้อมูล ภายใต้เงื่อนไขการใช้พื้นที่ในการจัดเก็บดัชนีที่จำกัด โดยการทำงานของระบบแบ่งออกเป็น 3 ขั้นตอน คือ 1. การคัดเลือกแอททริบิวต์ (Attribute) ที่ถูกสอบถามบ่อย และมีความเหมาะสมสำหรับการสร้างดัชนี Bitmap Join Index 2. การประเมินทรัพยากรที่ใช้ในการสร้างดัชนีสำหรับแอททริบิวต์แต่ละตัว 3. การเลือกดัชนีที่เหมาะสมที่สุดสำหรับพื้นที่ที่จำกัด โดยนำวิธีการแก้ปัญหาแบบพลวัต (Dynamic Programming) มาประยุกต์ใช้

คำสำคัญ : การสร้างดัชนีของคลังข้อมูล การคัดเลือกดัชนี วิธีแก้ปัญหาแบบพลวัต

Abstract

This paper presents Bitmap Join Index selection technique for improving query processing time in data warehouse environment under space usage constraints. The technique is consist of 3 steps i.e., Candidate Attribute Extraction, Resource Evaluation, and BJI Selection. Dynamic programming approach is used to find the optimal indices.

Keywords : Bitmap join index, Star schema, Index selection technique, Dynamic programming

1. บทนำ

คลังข้อมูล (Data Warehouse) เป็นฐานข้อมูลขนาดใหญ่ที่มีการเก็บรวบรวมข้อมูลที่มีความเกี่ยวข้องกันอย่างเป็นระบบ มีวัตถุประสงค์เพื่อนำมาใช้ในการวิเคราะห์ สรุปผล เพื่อสนับสนุนการตัดสินใจสำหรับผู้บริหารในองค์กรธุรกิจต่างๆ [2] ความต้องการพื้นฐานของผู้ใช้คลังข้อมูล ก็คือความรวดเร็วในการสอบถามข้อมูล [1] ด้วยเหตุนี้จึงจำเป็นต้องมีการเพิ่มประสิทธิภาพในการเข้าถึงข้อมูลเพื่อให้สามารถสอบถาม (Query) ได้รวดเร็ว ลักษณะโครงสร้างของคลังข้อมูลส่วนใหญ่เป็นแบบ Star Schema [2,8,12,13] ซึ่งประกอบด้วยตาราง Fact ซึ่งเป็นตารางหลักและเชื่อมโยงกับตาราง Dimension หลายตารางด้วยแอททริบิวต์ที่เป็นคีย์หลักของตาราง Dimension และหนึ่งในจำนวนตาราง Dimension ทั้งหมดต้องมีตารางที่เก็บข้อมูลเกี่ยวกับเวลา (Time Dimension) ดังภาพที่ 1(a) นอกจากนี้คำสั่งสอบถามข้อมูลสำหรับคลังข้อมูลส่วนใหญ่เป็นคำสั่งสอบถามที่มีความซับซ้อน (Complex Query) ซึ่งต้องใช้การ Join ตารางมากกว่า 1 ตาราง และการสอบถามเป็นลักษณะแบบทันทีทันใด (Ad hoc Query) ซึ่งประกอบด้วยเงื่อนไขการสอบถามข้อมูลที่ซับซ้อนขึ้นอยู่กับความต้องการของผู้ใช้ ทำให้ใช้เวลาในการสอบถามข้อมูลมาก [2,9]

การเพิ่มประสิทธิภาพการสอบถามข้อมูลในคลังข้อมูลนั้นมีหลายวิธี เช่น 1) การสร้าง Pre-computed Structure หรือ Summary Table และ 2) การสร้างดัชนี (Index) เป็นต้น

Pre-computed Structure คือ การสร้างตารางข้อมูลที่อาจถูกสอบถามไว้ล่วงหน้า ข้อดีคือ ช่วยเพิ่มความเร็วในการค้นหาข้อมูลได้มาก ถ้าคำสั่งที่สอบถามตรงกับที่สร้างไว้เพราะ

ไม่จำเป็นต้องอ่านข้อมูลจากตารางจริง ข้อเสียคือ เปลืองพื้นที่ในการเก็บข้อมูลมาก ไม่สามารถกำหนดได้ว่าควรสร้าง Pre-computed Structure สำหรับคำสั่งสอบถามใดบ้าง ดังนั้นจึงไม่เหมาะกับการสอบถามแบบทันทีทันใด และเมื่อใดที่มีการปรับปรุงข้อมูลในตารางจริง ก็ต้องมีการสร้าง Pre-computed Structure ขึ้นมาใหม่ทั้งหมด [2,10]

การสร้างดัชนี คือ วิธีการในการเพิ่มความรวดเร็วในการเข้าถึงข้อมูลโดยไม่ต้องมีการเพิ่มอุปกรณ์เข้ามาในระบบ ข้อดีคือ ช่วยให้ประหยัดค่าใช้จ่ายได้มาก และในคลังข้อมูลนั้นคำสั่งส่วนใหญ่เป็นคำสั่งสอบถาม (SELECT) และคำสั่งเพิ่มข้อมูล (INSERT) ซึ่งจะมีการเพิ่มข้อมูลเป็นช่วงเวลาที่น่าจะค่อนข้างหลายๆ [9] ดังนั้นการสร้างดัชนีจึงเป็นตัวเลือกที่ดีในการเพิ่มประสิทธิภาพการสอบถามของคลังข้อมูล ข้อเสียคือ หากมีการสร้างดัชนีจำนวนมาก จะช่วยลดเวลาในการสอบถามได้ดี แต่ต้องใช้พื้นที่เก็บดัชนีเป็นจำนวนมาก ซึ่งอาจมีดัชนีบางส่วนที่ไม่ได้ถูกเรียกใช้เลย ส่วนการสร้างดัชนีจำนวนน้อยเกินไป ก็จะทำให้ใช้เวลาในการเข้าถึงข้อมูลนาน

Bitmap Join Index เป็นดัชนีชนิดหนึ่งที่เป็นที่นิยมใช้เนื่องจากมีความเหมาะสมสำหรับคลังข้อมูล [3] ในบทความนี้ผู้วิจัยจึงได้เสนอแนวคิด และพัฒนาเทคนิคในการเลือกสร้างดัชนีสำหรับ Bitmap Join Index โดยคำนึงถึงทรัพยากรที่ใช้ในการสร้างดัชนีแต่ละตัว ว่าควรสร้างดัชนีสำหรับแอททริบิวต์ใดบ้างเพื่อให้เหมาะสมกับทรัพยากรที่จำกัด และยังคงมีประสิทธิภาพในด้านการสอบถามข้อมูล

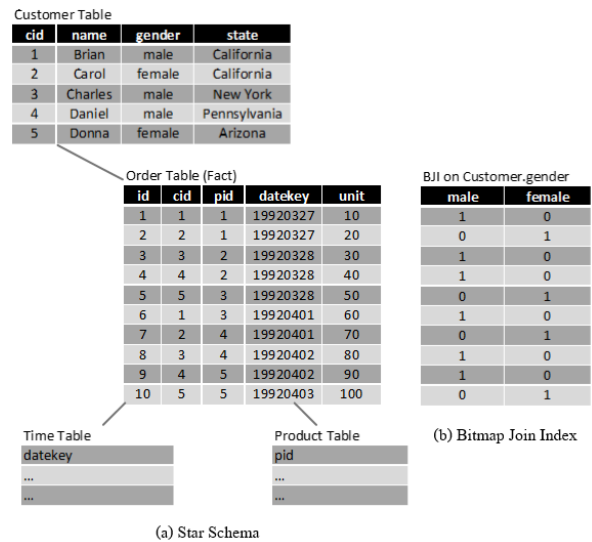
ในเอกสารนี้ประกอบด้วย 5 ส่วน โดยส่วนที่ 2 กล่าวถึงทฤษฎีและงานวิจัยที่เกี่ยวข้อง ส่วนที่ 3 กล่าวถึงขั้นตอนวิธีที่นำเสนอ ส่วนที่ 4 กล่าวถึงผลการดำเนินการวิจัย ส่วนที่ 5 กล่าวถึงบทสรุป และงานที่จะทำในอนาคต

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 Bitmap Join Index

ลักษณะของดัชนีแบบ Bitmap Join Index เป็นโครงสร้างตารางขนาด 2 มิติซึ่งแต่ละแถว (Row) เป็นรายการข้อมูลตาราง Fact และแต่ละคอลัมน์ (Column) เป็นค่าของแอททริบิวต์ที่ไม่ใช่คีย์หลักของตาราง Dimension [7] โดยค่าของแต่ละตำแหน่งที่เก็บในตารางจะเป็น 1 ก็ต่อเมื่อค่าข้อมูลจากตาราง

Dimension ตรงกับค่าข้อมูลจากตาราง Fact ส่วนตำแหน่งอื่นๆ จะเก็บค่า 0 ดังตัวอย่างในภาพที่ 1 (b) แสดง Bitmap Join Index บนตาราง Fact ซึ่ง Join กับตาราง Customer บนแอททริบิวต์เพศ



ภาพที่ 1: ลักษณะของ Bitmap Join Index

สำหรับการสอบถาม โดยใช้ดัชนี Bitmap Join Index นั้นสามารถใช้การดำเนินการทางตรรกะ AND OR และ NOT เข้ามาใช้เมื่อคำสั่งสอบถามมีการสอบถามข้อมูลจากหลายตาราง และมีเงื่อนไขในการสอบถามมากกว่า 1 แอททริบิวต์ [9,13]

2.2 วิธีการกำหนดการพลวัต (Dynamic Programming)

เป็นการแบ่งปัญหาใหญ่ที่ซับซ้อนออกเป็นปัญหาย่อยๆ แล้วหาผลลัพธ์ของปัญหาย่อยนั้น จากนั้นจะมีการเก็บผลลัพธ์เอาไว้สำหรับการเรียกใช้เมื่อต้องการแก้ปัญหานั้นที่ใหญ่ขึ้น หรือมีการเรียกใช้ซ้ำโดยไม่ต้องมีการคำนวณใหม่ ตัวอย่างเช่น ปัญหาการคำนวณหาเส้นทางที่สั้นที่สุด All-Pairs Shortest-Paths Problem และ ปัญหาการเลือกวัตถุที่เหมาะสมที่สุด Knapsack Problem เป็นต้น [14]

2.3 งานวิจัยที่เกี่ยวข้อง

ปัญหาการเลือกดัชนีเป็นปัญหาที่มีความท้าทายเป็นอย่างมาก เนื่องจากในการคัดเลือกดัชนีจะมีจำนวนแอททริบิวต์ที่สามารถถูกเลือกได้เป็นจำนวนมาก แต่เนื่องจากพื้นที่ที่ใช้ในการเก็บข้อมูลมีอยู่อย่างจำกัด จึงไม่สามารถสร้างดัชนีสำหรับการสอบถามได้ งานวิจัยที่ถูกคิดค้นมาเพื่อแก้ปัญหานี้ เช่น Aouiche และคณะ [4] ได้นำเสนอวิธีการในการเลือกดัชนี

โดยการนำเอาเทคนิคในการหากลุ่มข้อมูลที่ปรากฏด้วยกันบ่อยมาใช้ในการคัดเลือก Candidate Index หลังจากนั้นมีการสร้าง Cost Model ขึ้นมาและใช้เทคนิค Greedy ในการคัดเลือกดัชนีที่มีความคุ้มค่ามากที่สุด สำหรับเงื่อนไขของพื้นที่ที่จำกัด

ในปี 2007 Bellatreche และคณะ [5] ได้พัฒนาวิธีการสำหรับการแก้ปัญหาการเลือกดัชนี โดยมีการเพิ่มการพิจารณาปัจจัยที่เกี่ยวข้องกับการสร้างดัชนี ในการคัดเลือกแอททริบิวต์ จากนั้นใช้เทคนิค Greedy ในการเลือกแอททริบิวต์ที่จะนำไปสร้างดัชนี

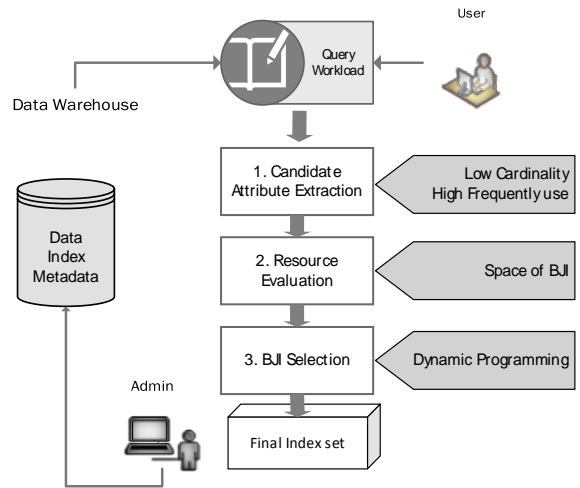
ในปี 2011 Ziani และ Ouinten [6] ใช้วิธีการเลือกเฉพาะชุดแอททริบิวต์ที่ไม่เป็นเซตย่อยของกลุ่มข้อมูลอื่น เพื่อลดจำนวน Candidate Index ที่จะนำมาสร้างดัชนี และทำการตัดกลุ่มข้อมูลที่ไม่จำเป็นออกไปเพื่อลดจำนวนแอททริบิวต์ที่จะทำการคัดเลือก

ในปี 2012 Hyoung และ Jae [7] ได้นำเสนอวิธีเลือกดัชนี โดยแบ่งการทำงานเป็น 2 ขั้นตอนหลักคือ การลดจำนวน Candidate Index โดยพัฒนาอัลกอริทึม EFP-Tree ซึ่งได้ปรับปรุงจากเทคนิค FP-Growth โดยการกำจัดกลุ่มข้อมูลที่มีความเกี่ยวข้องกับแอททริบิวต์ที่เป็นคีย์หลักออกไป และปรับปรุงอัลกอริทึม sBJI เพื่อใช้ในการคัดเลือกดัชนีจาก Candidate Itemsets ภายใต้งานวิจัยที่เกี่ยวข้องทั้งหมดยังไม่ได้มีการใช้วิธีการการแก้ปัญหาแบบพลวัตซึ่งเป็นวิธีการที่มีประสิทธิภาพ และเมื่อต้องการปรับเปลี่ยนขนาดของพื้นที่ที่ใช้ก็สามารถได้ผลลัพธ์ที่ต้องการทันที เนื่องจากมีการเก็บข้อมูลผลลัพธ์เอาไว้แล้ว

งานวิจัยที่เกี่ยวข้องทั้งหมดยังไม่ได้มีการใช้วิธีการการแก้ปัญหาแบบพลวัตซึ่งเป็นวิธีการที่มีประสิทธิภาพ และเมื่อต้องการปรับเปลี่ยนขนาดของพื้นที่ที่ใช้ก็สามารถได้ผลลัพธ์ที่ต้องการทันที เนื่องจากมีการเก็บข้อมูลผลลัพธ์เอาไว้แล้ว

3. Bitmap Join Index Selection (BJIS) สำหรับคลังข้อมูล

ขั้นตอนวิธีการทำงานของ BJIS ประกอบไปด้วยขั้นตอนหลักทั้งหมด 3 ขั้นตอน คือ 1.การคัดเลือกแอททริบิวต์ที่ถูกสอบถามบ่อย และมีความเหมาะสมสำหรับการสร้างดัชนี Bitmap Join Index (Candidate Attribute Extraction) 2. การประเมินทรัพยากรที่ใช้ในการสร้างดัชนีสำหรับแอททริบิวต์แต่ละตัว (Resource Evaluation) และ 3.การเลือกดัชนีที่เหมาะสมที่สุดสำหรับพื้นที่ที่จำกัด (BJI Selection) ดังแสดงในภาพที่ 2



ภาพที่ 2: โครงสร้างการทำงานของระบบ BJIS

สามารถอธิบายขั้นตอนการทำงานได้ดังต่อไปนี้

3.1 Candidate Attribute Extraction

เป็นขั้นตอนในการคัดเลือก Candidate Attribute ซึ่งแบ่งเป็นขั้นตอนย่อย 2 ขั้นตอนดังนี้

1) สกัดแอททริบิวต์ทั้งหมดจากชุดคำสั่งสอบถามที่ผ่านมาในอดีต (Workload) และคัดเลือกแอททริบิวต์ที่มีความเหมาะสมสำหรับการสร้างดัชนีแบบ Bitmap Join Index นั่นคือ เป็นแอททริบิวต์ที่มีคาร์ดินอลลีต่ำ (Low Cardinality)

2) คัดเลือกเฉพาะแอททริบิวต์ที่มีการสอบถามบ่อย โดยการหาค่าความถี่ในการถูกสอบถามจากชุดข้อมูลการสอบถามที่ผ่านมาในอดีต ซึ่งจะเลือกเฉพาะแอททริบิวต์ที่มีความถี่ในการสอบถามสูงกว่าค่าความถี่ขั้นต่ำ (Minimum Frequent) ซึ่งกำหนดโดยผู้ดูแลระบบ

3.2 Resource Evaluation

ในขั้นตอนนี้ระบบจะคำนวณหาทรัพยากรที่ใช้ในการสร้างดัชนี (เช่น พื้นที่ที่ใช้ในการเก็บดัชนี) ของ Candidate Attribute แต่ละตัว ในขั้นตอนนี้ Candidate Attributes Table (CAT) จะถูกสร้างขึ้น ในที่นี้พื้นที่ที่ใช้ในการเก็บดัชนีสามารถคำนวณได้จากจำนวนคาร์ดินอลลีของแต่ละแอททริบิวต์คูณกับจำนวนแถวในตาราง Fact ซึ่งผลลัพธ์ที่ได้จากขั้นตอนนี้จะถูกเก็บใน CAT เพื่อนำไปใช้ในขั้นตอน BJI Selection

ขั้นตอนวิธี BJI Selection

ข้อมูลนำเข้า : - ขนาดของพื้นที่ทั้งหมด , C

- พื้นที่จัดเก็บดัชนีของแอททริบิวต์ i, W_i
- ค่าความถี่ของแอททริบิวต์ i, V_i
- จำนวน Candidate Attributes ทั้งหมด, n

ข้อมูลส่งออก : - ชุดดัชนีที่ถูกเลือกภายในพื้นที่ C

1. กำหนดค่าเริ่มต้นให้กับตาราง $V[0..n,0..C]$
 - 1.1 กำหนดให้ทุกค่าในแถวที่ 0 เท่ากับ 0
 - 1.2 กำหนดให้ทุกค่าในคอลัมน์ที่ 0 เท่ากับ 0
 - 1.3 กำหนดให้ค่าในตาราง $(V[1..n,1..C])$ เท่ากับ -1
 - 1.4 กำหนดค่าเริ่มต้นให้ $i=n$ และ $j=C$
2. BJI_Selection(i,j)
 - 2.1 if ($V[i][j]<0$)
 - if ($j<W[i]$)
 - $V[i][j]=$ BJI_Selection ($i-1,j$);
 - else
 - $V[i][j]=\max[$ BJI_Selection ($i-1,j$),
 $V[i]+$ BJI_Selection ($i-1,j-W[i]$)];
 - 2.2 เก็บค่า $V[i][j]$ เป็นมูลค่าที่สูงที่สุดที่ตำแหน่ง (i,j)
3. FindIndexSet(i,j)
 - 3.1 For ($i=n,j=C ; j>0 \ \&\& \ i>0 ,i--$)
 - if ($V[i][j] \neq V[i-1][j]$)
 - เก็บค่าแอททริบิวต์ที่ i ในชุดดัชนีที่ถูกเลือก
 - $j=j-W[i]$;
 - 3.2 Return Final Index Set

ภาพที่ 3: ขั้นตอนวิธีในการเลือกดัชนี

3.3 BJI Selection

ในขั้นตอนนี้จะเป็นการเลือกดัชนีที่เหมาะสมที่สุดภายใต้ทรัพยากรที่จำกัด โดยเทคนิคที่เลือกมาประยุกต์ใช้ได้แก่ 0-1 Knapsack [14] ซึ่งปัจจัยที่ใช้ในการพิจารณาได้แก่ ขนาดของพื้นที่ที่ใช้ในการเก็บดัชนี และ ค่าความถี่ของแต่ละแอททริบิวต์ โดยใช้ขั้นตอนวิธีดังแสดงในภาพที่ 3 สำหรับผลลัพธ์ที่ได้จะเป็นชุดข้อมูลแอททริบิวต์ที่ถูกเลือกเพื่อนำไปสร้างดัชนี Bitmap Join Index

การทำงานของขั้นตอนวิธี BJI Selection แบ่งเป็น 2 ส่วน คือ ส่วนของฟังก์ชัน BJI_Selection (i,j) ซึ่งเป็นขั้นตอนในการหามูลค่าที่มากที่สุดที่เป็นไปได้เมื่อมีจำนวนแอททริบิวต์เท่ากับ i และพื้นที่ที่ใช้ในการจัดเก็บดัชนี j และส่วนของ

ฟังก์ชัน FindIndexSet (i,j) เป็นขั้นตอนวิธีในการหาชุดคำตอบของ แอททริบิวต์ที่ถูกเลือกทั้งหมด

ตัวอย่างการทำงานของ BJIS

ขั้นตอนที่ 1 Candidate Attributes Extraction สกัดแอททริบิวต์ที่จะนำมาสร้างดัชนีจากชุดคำสั่งสอบถาม ซึ่งมีลักษณะดังภาพที่ 4 โดยคัดเลือกแอททริบิวต์ที่มีคาร์ดินอลิตี้ต่ำ จากนั้นคำนวณหาค่าความถี่ของการสอบถามของแต่ละแอททริบิวต์ เมื่อกำหนดให้ค่าสนับสนุนขั้นต่ำเท่ากับ 40 % จะได้ชุดแอททริบิวต์ผลลัพธ์คือ {a,b,c,d,e,f} ส่วนแอททริบิวต์ {g,h} มีค่าสนับสนุนน้อยกว่าค่าสนับสนุนขั้นต่ำจึงไม่ถูกคัดเลือกดังแสดงในตารางที่ 1

```
SELECT *
FROM Fact F, Dimension1 D1, Dimension2 D2
WHERE AND F.D1id=D1.D1id AND
F.D2id=D2.D2id AND D1.a=1 AND D1.b=2
AND D2.c=3 AND D2.e=4
```

ภาพที่ 4: ตัวอย่างชุดคำสั่งสอบถาม

ตารางที่ 1: แอททริบิวต์ที่ปรากฏหลังเงื่อนไขการสอบถาม

คำสั่งสอบถาม	แอททริบิวต์ทั้งหมดที่ปรากฏหลังเงื่อนไขการสอบถาม
1	a, b, c, e
2	b, c, d, g
3	c, d, h
4	a, b, e, f
5	b, e, f

ขั้นตอนที่ 2 Resource Evaluation เมื่อได้จำนวน

Candidate Attributes ทั้งหมดแล้ว ก็จะคำนวณหาพื้นที่ที่ใช้ในการสร้างดัชนีสำหรับแอททริบิวต์แต่ละตัว เพื่อนำไปใช้ในขั้นตอนเลือกดัชนี BJI Selection จะได้ผลลัพธ์ดังตารางที่ 2

ตารางที่ 2: Candidate Attributes Table (CAT)

แอททริบิวต์	พื้นที่ที่ใช้	ความถี่
a	2	2
b	4	4
c	1	3
d	3	2
e	6	3
f	5	2

ขั้นตอนที่ 3 BJI Selection นำข้อมูลจาก CAT มาดำเนินการตามขั้นตอนวิธีการเลือกดัชนีในภาพที่ 3 ซึ่งผลลัพธ์ที่ได้จากการทำงานของกระบวนการนี้แสดงตารางที่ 3 (ใช้ $n=6, C=6$)

ตารางที่ 3: ผลลัพธ์ที่ได้จากการทำงานของขั้นตอนวิธีการเลือกดัชนี

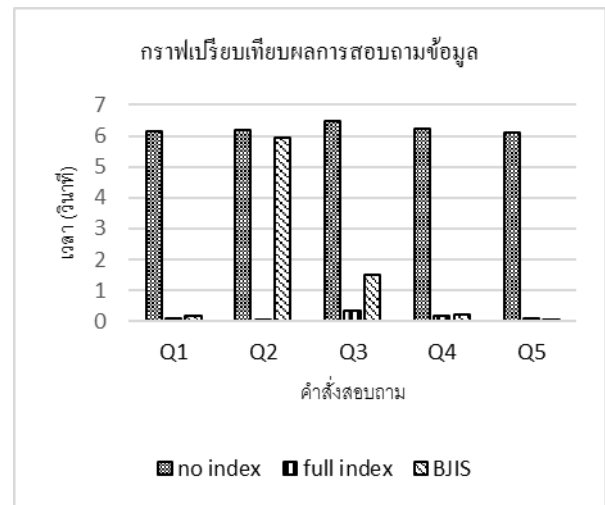
			$\begin{matrix} i \\ \backslash \\ i \end{matrix}$	0	1	2	3	4	5	6
Attribute	w_i	v_i	0	0	0	0	0	0	0	0
c	1	3	1	0	3	3	3	3	3	3
a	2	2	2	0	3	3	5	5	5	5
b	4	4	3	0	3	3	5	5	7	7
d	3	2	4	0	3	3	5	5	7	7
e	6	3	5	0	3	3	5	5	7	7
f	5	2	6	0	3	3	5	5	7	7

เมื่อพิจารณาผลลัพธ์ที่ได้จากขั้นตอนวิธี BJI Selection ในบางครั้งอาจพบว่ามีชุดข้อมูลอื่นอีกที่ให้มูลค่าเท่ากับชุดข้อมูลที่ถูกเลือก (เนื่องจากแอททริบิวต์ที่พบก่อนจะถูกเลือกก่อนเสมอ) เช่น ดังข้อมูลในตารางที่ 3 เมื่อกำหนดขนาดพื้นที่ที่ใช้เท่ากับ 6 จะเห็นได้ว่าชุดข้อมูลที่เป็นคำตอบสามารถมีได้มากกว่า 1 ชุด ที่ให้มูลค่าที่สูงที่สุดเท่ากัน คือ 7 ได้แก่ {c,b} และ {c,a,d} ซึ่งใช้พื้นที่รวมเท่ากับ 5 และ 6 ตามลำดับ ซึ่งยังเป็นไปตามเงื่อนไขของการใช้พื้นที่ที่เหมาะสมแต่ผลลัพธ์ที่ดีที่สุดคือชุดข้อมูลชุด {c,b} เนื่องจากมีการใช้พื้นที่น้อยกว่าชุด {c,a,d} ดังนั้นหากมีการเรียงลำดับข้อมูลตามความคุ้มค่าของข้อมูล (ความถี่/พื้นที่) แต่ละแอททริบิวต์ของข้อมูลจากมากไปน้อย ก่อนที่จะนำไปคำนวณตามวิธีการ BJI Selection จะได้ผลลัพธ์ที่เป็นชุดข้อมูลที่มีมูลค่ามากที่สุด และมีการใช้พื้นที่น้อยที่สุด

4. ผลการดำเนินการวิจัย

การทดสอบประสิทธิภาพของขั้นตอนวิธี BJIS นั้น ผู้วิจัยได้ทดสอบโดยการคัดเลือกชุดคำสั่งตัวอย่างจากคำสั่งสอบถามทั้งหมดของ SSB Benchmark [11] ซึ่งเป็นตัวอย่างคลังข้อมูลที่มีลักษณะโครงสร้างแบบ Star Schema โดยเลือกคำสั่งสอบถามข้อมูล RDBMS Q3.2(Q1), Q3.4(Q2), Q4.1(Q3), Q4.2(Q4) และ Q4.3(Q5) สำหรับการทดสอบประสิทธิภาพของระบบจะเปรียบเทียบผลการทดลองเป็น 3 กลุ่ม คือ กลุ่มที่

ไม่มีการสร้างดัชนีใดเลย(no index) กลุ่มที่มีการสร้างดัชนีสำหรับทุกแอททริบิวต์(full index) และกลุ่มที่สร้างดัชนีโดยใช้ขั้นตอนวิธี BJIS ได้ผลลัพธ์ดังที่ปรากฏในภาพที่ 5



ภาพที่ 5 : ผลการทดสอบขั้นตอนวิธี BJIS

จากผลการทดลองจะเห็นได้ว่าการเลือกสร้างดัชนีโดยใช้ขั้นตอนวิธี BJIS นั้น คำสั่งสอบถาม Q1, Q3, Q4 และ Q5 จะใช้เวลาในการสอบถามใกล้เคียงกับการสอบถามข้อมูลเมื่อมีการสร้างดัชนีสำหรับทุกแอททริบิวต์ สำหรับคำสั่งสอบถามที่ Q2 ใช้เวลาในการสอบถามใกล้เคียงกับการสอบถามเมื่อมีการสร้างดัชนีสำหรับทุกแอททริบิวต์ เนื่องจากชุดดัชนีที่ถูกเลือกนั้นไม่ตรงกับแอททริบิวต์ที่ถูกสอบถามในคำสั่งสอบถาม Q2 จึงไม่ได้ช่วยเพิ่มประสิทธิภาพในการสอบถาม เมื่อพิจารณาอัตราส่วนของพื้นที่ที่ใช้ในการสร้างดัชนีของทั้ง 3 กลุ่ม (no index : full index : BJIS) เท่ากับ 0 : 100 : 38

ดังนั้นสามารถสรุปได้ว่า การเลือกสร้างดัชนี Bitmap Join Index โดยใช้ขั้นตอนวิธี BJIS นั้น ทำให้สามารถเลือกดัชนีได้อย่างถูกต้อง เมื่อระบบมีพื้นที่ในการสร้างดัชนีที่จำกัด เนื่องจากสามารถลดเวลาในการสอบถามข้อมูลได้เป็นอย่างดี

5. สรุปผลและงานที่จะทำในอนาคต

ในการสอบถามข้อมูลจากคลังข้อมูลนั้น มีความจำเป็นอย่างยิ่งที่จะต้องมีการเพิ่มประสิทธิภาพในการสอบถามเนื่องจากการเข้าถึงข้อมูลขนาดใหญ่ และคำสั่งสอบถามที่มีความซับซ้อนนั้น จะทำให้ใช้เวลาในการประมวลผลมาก ในบทความนี้ได้นำเสนอระบบแนะนำการคัดเลือกดัชนี Bitmap Join Index ซึ่ง

ช่วยให้ผู้ใช้สามารถเลือกสร้างดัชนีได้อย่างมีประสิทธิภาพภายใต้ทรัพยากรที่จำกัด โดยมีขั้นตอนในการเลือกแอททริบิวต์ที่มีโอกาสถูกสอบถามบ่อยเพื่อนำมาสร้างดัชนี และใช้เทคนิคการแก้ปัญหาแบบพลวัต ซึ่งเหมาะสำหรับการแก้ปัญหาที่มีความซับซ้อน ทำให้ได้ผลลัพธ์ที่เหมาะสมที่สุด และมีความยืดหยุ่นเมื่อต้องการลดขนาดของพื้นที่ที่ใช่ เนื่องจากมีการเก็บผลลัพธ์ของปัญหาย่อยเอาไว้แล้ว

สำหรับงานที่จะทำในอนาคตได้แก่ การนำวิธีการสร้างดัชนีบิตแมปชนิดต่างๆ มาใช้ในการลงรหัส เช่น ดัชนีบิตแมปแบบเข้ารหัส ดัชนีบิตแมปแบบกระจาย เป็นต้น เพื่อช่วยลดขนาดของพื้นที่ที่ใช้ในการจัดเก็บข้อมูล และคิดปัจจัยอื่นๆ ที่มีผลต่อทรัพยากรที่ใช้ในการสร้างดัชนี

เอกสารอ้างอิง

- [1] M. Golfarelli, S. Rizzi and E. Saltarelli, "Index selection for data warehousing," *Proc. DMDW*, pp. 33-42, 2002.
- [2] S. Chaudhuri and U. Dayal, "An overview of data warehousing and OLAP technology," *ACM Sigmod record 26.1*, pp. 65-74, 1997.
- [3] M. C. Wu and A. P. Buchmann, "Research issues in data warehousing," *Datenbanksysteme in Büro, Technik und Wissenschaft.*, pp. 61-82, 1997.
- [4] K. Aouiche, J. Darmont, O. Boussaïd and F. Bentayeb, "Automatic selection of bitmap join indexes in data warehouses," *Data Warehousing and Knowledge Discovery.*, pp. 64-73, 2005.
- [5] L. Bellatrecheet, R. Missaoui, H.Necir and H. Drias, "A Data Mining Approach for selecting Bitmap Join Indices," *JCSE 1.2*, pp. 177-194, 2007.
- [6] B. Ziani and Y. Ouinten, "Improving Star Join Queries Performance: A Maximal Frequent Pattern Based Approach for Automatic Selection of Indexes in Relational Data Warehouses," *Internet Computing & Information Services (ICICIS), 2011 International Conference on. IEEE*, 2011.
- [7] H. G. An and J. J. Koh, "A Study on the Selection of Bitmap Join Index using Data Mining Techniques," *Strategic Technology (IFOST), 7th International Forum on. IEEE*, pp. 1-5, 2012.
- [8] H. Gupta, V. Harinarayan, A. Rajaraman and J. D. Ullman, "Index selection for OLAP," *Proceedings. 13th International Conference on. IEEE*, p. 208-219, 1997.
- [9] P. O'Neil and D. Quass, "Improved query performance with variant indexes," *ACM Sigmod Record.*, Vol. 26, No. 2, pp.38-49, 1997.
- [10] P. Bizarro and H. Madeira, "The Dimension-Join: A New Index for Data Warehouses.", *SBBD*, pp. 259-273, 2001.
- [11] P. O'Neil, E. O'Neil and X. Chen, "The star schema benchmark," [online] 2009. Availablefrom: <http://www.cs.umb.edu/~poneil/StarSchemaB.PDF> [2013, September 10].
- [12] R. Kimball, *The data warehouse toolkit.*, John Wiley & Sons, Inc., 2006.
- [13] L. Bellatrecheet, R. Missaoui, H.Necir and H. Drias, "Selection and pruning algorithms for bitmap index selection problem using data mining," *Data Warehousing and Knowledge Discovery.*, pp.221-230, 2007.
- [14] A. Levitin, *Introduction to the design & analysis of algorithms*, Pearson Addison-Wesley, 2007.

เทคนิคความจริงเสริมเพื่อนำเสนอข้อมูลแบบบ้านบ้านจัดสรร

Applied AR Technique to Present Information of Real Estate House Plans

สันติส จุลโพธิ์ (Suntis Jullapho)¹ และณัฐวี ฤกษ์ฤกษ์ (Nattavee Utakrit)²

^{1,2}ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

suntis.j@gmail.com¹, nattaveeu@kmutnb.ac.th²

บทคัดย่อ

การประยุกต์ใช้เทคนิคความจริงเสริมในธุรกิจอสังหาริมทรัพย์ประเภทบ้านจัดสรรในปัจจุบันยังมีไม่มากนัก งานวิจัยนี้จึงได้ทำการศึกษาและประยุกต์ใช้เทคนิคความจริงเสริมเพื่อนำเสนอข้อมูลแบบบ้านจัดสรร โดยการศึกษาครั้งนี้เลือกใช้กรณีศึกษาจาก โครงการหมู่บ้านของ บริษัท กรุงเทพบ้านและที่ดิน จำกัด (มหาชน) โดยมีแบบบ้านทั้งหมด 5 แบบในการทดลอง การนำเสนอข้อมูลประเภทดังกล่าวนี้ ผู้วิจัยได้พัฒนาเป็นรูปแบบการทำงานแบบผสมผสานกันระหว่างเทคนิคความจริงเสริมร่วมกับเว็บแอปพลิเคชันที่พัฒนาขึ้นเพื่อช่วยให้นำเสนอรูปแบบการนำเสนอสินค้า มีความสมบูรณ์ และอำนวยความสะดวกให้พนักงานขายสามารถนำเสนอข้อมูลการขายบ้านได้ในรูปแบบที่แปลกใหม่ และช่วยลูกค้าสามารถมองเห็นแผนผังภาพรวมโครงการ รวมไปถึงโมเดลแบบบ้านเสมือนจริง 3 มิติ ทั้งนี้เพื่อเป็นข้อมูลในการตัดสินใจได้ดียิ่งขึ้นอีกด้วย และเป็นการสร้างความได้เปรียบในการแข่งขัน

ผลการวิจัยพบว่า เมื่อนำระบบไปทดลองใช้งาน ระบบสร้างจุดเด่นและความน่าสนใจให้กับผู้ใช้งานเป็นอย่างมาก โดยผลการประเมินความพึงพอใจอยู่ในระดับดีมาก รวมไปถึงการทดสอบประสิทธิภาพการทำงานมีค่าถึงร้อยละ 97.56 จัดอยู่ในระดับดีมากเช่นกัน

คำสำคัญ: ความจริงเสริม อสังหาริมทรัพย์ บ้านจัดสรร

Abstract

Currently, there are only a few augmented reality techniques in the field of real estate. This new augmented reality research offers comprehensive education on how to apply and present information about these real estate techniques. This research uses data compiled from Krungthep

Land Public Company Limited and consists of 5 house plans. This system blends augmented reality and web-application development technologies together. That way, sales team can present this data to the customer in an easy to understand and user-friendly way. Customers can view the overall project plan and virtual 3D house. When the customer selects the product, the system offers house-pricing and status of sale. This allows the customer to see every option before construction begins, including a 3D to-scale version of how the house will look upon completion to ensure total satisfaction during their decision-making process.

The responses has been good, dominant, and attractive. The system was evaluated by user's satisfaction and performance test. The results of the satisfaction is great and test performance of 97.56% is very good as well.

Keyword: Augmented Reality, Real Estate, Housing Development.

1. บทนำ

ปัจจุบันธุรกิจอสังหาริมทรัพย์เป็นธุรกิจที่มีการแข่งขันกันค่อนข้างสูง การเปิดตัวโครงการใหม่เพิ่มขึ้นอย่างต่อเนื่อง ผู้ประกอบการต่างทำแคมเปญการตลาดด้านต่าง ๆ เพื่อกระตุ้นให้ผู้บริโภคตัดสินใจซื้อที่อยู่อาศัยเร็วขึ้น การนำเสนอข้อมูลการขายมีมากมายหลายวิธี การสร้างความได้เปรียบทางการแข่งขันเป็นเรื่องที่สำคัญ ผู้ประกอบการจะต้องหาวิธีนำเสนอข้อมูลให้มีความน่าสนใจ โดยเทคนิคความจริงเสริม (Augmented Reality) เป็นรูปแบบหนึ่งที่น่าสนใจนำมาประยุกต์ใช้งาน

แนวโน้มการนำเทคนิคความจริงเสริมมาประยุกต์ใช้งานมีมากขึ้นเป็นลำดับในหลากหลายรูปแบบ [1] ตัวอย่างเช่นหน่วยงานกรมอสังหาริมทรัพย์ ประเทศไทย ได้นำมาใช้รณรงค์ความ

ปลอดภัยบนท้องถนน ตามจุดที่เกิดอุบัติเหตุบ่อย ๆ โดยติดตั้ง QR Code ไว้เพื่อให้ผู้ที่ผ่านไปมาสงสัยแล้วสแกนเพื่อทดสอบ ซึ่งจะแสดงสัญลักษณ์ลอยขึ้นมาบนหน้าจอแสดงสาเหตุและสถิติของอุบัติเหตุ [2]

จากรายงานการวิจัยของ Jupiter Research ในปี 2013 [3] พบว่ามีผู้ใช้งานแอปพลิเคชันเทคนิคความจริงเสริมเป็นประจำประมาณ 60 ล้านคน และจะเพิ่มเป็น 200 ล้านคนในปี 2018 นวัตกรรมของเทคนิคความจริงเสริมเริ่มต้นมาจากการพัฒนาเพื่อใช้งานกับเกมส์ วิถีชีวิต และการออกกำลังกาย โดยในปี 2018 การใช้งานเทคนิคความจริงเสริมร่วมกับเกมส์และการใช้ตำแหน่งทางภูมิศาสตร์จะลดสัดส่วนลงกว่าครึ่งหนึ่งของตลาด โดย Jupiter Research เชื่อว่าจะมีการใช้งานด้านการดูแลสุขภาพ วิถีการดำเนินชีวิต การซื้อขายสินค้า และอากาศยานเพิ่มมากขึ้น

จากแนวโน้มการใช้งานที่เพิ่มมากขึ้น แต่อย่างไรก็ตามการนำเทคนิคดังกล่าวมาใช้ในธุรกิจประเภทอสังหาริมทรัพย์ยังมีไม่มากนัก ดังนั้นเพื่อสร้างความได้เปรียบในการแข่งขัน และสร้างความแตกต่างเหนือคู่แข่ง ผู้วิจัยจึงศึกษาเทคนิคความจริงเสริมและนำมาประยุกต์ใช้ เพราะเป็นทางเลือกที่น่าสนใจสำหรับการนำมาใช้แสดงโมเดลบ้าน ซึ่งทำให้ลูกค้าเกิดความแปลกใหม่ในรูปแบบการแนะนำโครงการ และเป็นตัวช่วยพนักงานขายของโครงการในการขายบ้านจัดสรรได้ดียิ่งขึ้น อีกทั้งลูกค้ายังสามารถมองเห็นภาพ 3 มิติ เสมือนจริงของโมเดลบ้านอีกด้วย

2. งานวิจัยที่เกี่ยวข้อง

สุพรรณพงษ์ [4] ได้ประยุกต์ใช้เทคนิคความจริงเสริมเพื่อใช้ในการสอนเรื่องพยัญชนะภาษาไทย เนื่องจากการพัฒนาระบบดังกล่าวจะช่วยให้อาจารย์ผู้สอนมีรูปแบบในการเรียนการสอนที่แปลกใหม่ ช่วยเพิ่มความน่าสนใจแก่นักเรียน ส่วน มงคล [5] พัฒนาระบบสารสนเทศด้านการท่องเที่ยวด้วยความเป็นจริงเสริมบนโทรศัพท์เคลื่อนที่ โดยอ้างอิงตามตำแหน่งตามพื้นที่ของผู้ใช้งานผสมผสานกับเทคนิคความจริงเสริมให้แก่นักท่องเที่ยวเข้าถึงข้อมูลต่าง ๆ เกี่ยวกับการท่องเที่ยวได้ทุกที่ทุก

ณัฐพล [6] ได้นำระบบความจริงเสริมสำหรับการถ่ายทอดทักษะการประกอบชิ้นงาน โดยเพิ่มเติมข้อมูลให้กับผู้ใช้ ซึ่งผู้ที่เข้ารับการถ่ายทอดทักษะจะมองเห็นสภาพแวดล้อมในการทำงานจริงกับกราฟิกเสริมได้ในหน้าการแสดงผลเดียวกันและ

พร้อมกัน และ พีรศิลป์ [7] พัฒนาระบบความจริงเสริมบนสภาพแวดล้อมที่เน้นการสัมผัสสำหรับการเรียนรู้ เป็นการประยุกต์ใช้สำหรับการเรียนรู้เรื่องการสังเคราะห์เสียง โดยควบคุมการทำงานผ่านการสัมผัสแบบหลายจุดบนพื้นผิวของโต๊ะได้ 2 ลักษณะคือ ใช้นิ้วกดเพื่อเลือก และใช้นิ้วลากเพื่อปรับเปลี่ยนค่า

Henderson และ Feiner [8] พัฒนาในหัวข้อ Augmented Reality for Maintenance and Repair เป็นการประยุกต์ใช้เทคนิคความจริงเสริมในการแสดงบำรุงรักษาและซ่อมอุปกรณ์ โดยใช้จอภาพแบบสวมศีรษะ (Head Mounted Display) ระบบจะแสดงรายละเอียดของชิ้นส่วนต่าง ๆ ของอุปกรณ์ที่ทำการซ่อมพร้อมชี้ตำแหน่ง และบอกวิธีการแก้ไขเป็นลำดับขั้นตอน

จากการศึกษาวิจัยที่เกี่ยวข้องทำให้ผู้วิจัยพบว่า ปัญหาส่วนใหญ่ที่พบในการทำการวิจัยของแต่ละงานคือ ประสิทธิภาพของเครื่องที่ใช้ในการวิจัย การตรวจจับแผ่นสัญลักษณ์ (Marker) กล้องเว็บแคมมีประสิทธิภาพต่ำเกินไป อัลกอริทึมที่ซับซ้อนเกินไปส่งผลให้การทำงานในภาพรวมสะดุด เกิดปัญหาในการประมวลผล บางงานวิจัยรูปแบบการนำเสนอไม่หลากหลาย เข้าใจยาก อุปกรณ์ที่นำมาใช้มีประสิทธิภาพไม่เพียงพอแต่บางอุปกรณ์ก็ไม่สามารถนำที่มีประสิทธิภาพสูงมาใช้งานได้เนื่องจากมีราคาสูง การรองรับอุปกรณ์ที่ใช้งานยังไม่หลากหลาย ผู้วิจัยจึงนำข้อคิดและปัญหาเหล่านี้มาหาทางแก้ไขปรับปรุง คัดเลือกเทคโนโลยีและเครื่องมือที่เหมาะสมมาใช้งาน โดยงานวิจัยนี้ได้นำเครื่องมือ AR-media™ มาประยุกต์ใช้ ซึ่งผู้ใช้งานสามารถใช้งานได้หลากหลายระบบปฏิบัติการ ได้แก่ Windows XP/Vista/7/8, Mac OS X 10.5 and later (Intel), iOS และ Android ช่วยให้มีทางเลือกในการใช้งาน อีกทั้งต้องออกแบบให้การใช้งานง่ายและน่าสนใจ อันจะส่งผลให้สามารถเพิ่มประสิทธิภาพ และต่อ ยอดการดำเนินธุรกิจที่เกี่ยวข้องต่อไป

3. วิธีการดำเนินการวิจัย

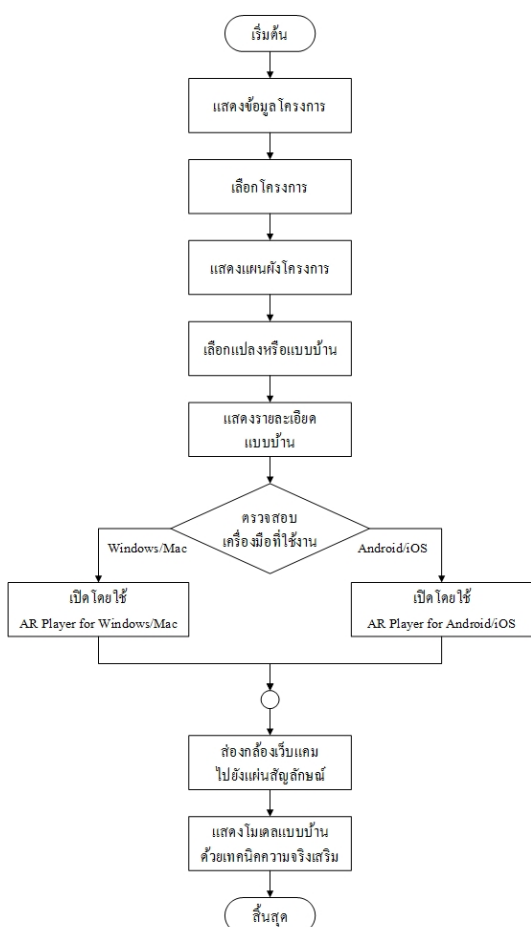
3.1 การศึกษาและวิเคราะห์ข้อมูล

ในขั้นตอนนี้ต้องศึกษารูปแบบและเทคโนโลยีที่ใช้ในการแสดงผลข้อมูลที่น่าสนใจ หลักการและการทำงานของเทคนิคความจริงเสริมจากเอกสารที่เกี่ยวข้องและผลงานวิจัย รวมไปถึงเครื่องมือต่าง ๆ ที่จำเป็นในการพัฒนาระบบ จากนั้นมาวิเคราะห์ถึงข้อดีข้อเสียของแต่ละรูปแบบแล้วนำมาปรับใช้งานต่อไป

3.2 การออกแบบระบบ

การออกแบบระบบเพื่อพัฒนาระบบช่วยนำเสนอข้อมูลแบบบ้านจัดสรรด้วยเทคนิคความจริงเสริมนี้ เป็นขั้นตอนที่สำคัญมาก ซึ่งจะทำให้การพัฒนาระบบมีประสิทธิภาพ โดยมีรายละเอียดดังนี้

แผนผังการทำงานของผู้ใช้งานระบบ เริ่มจากหน้าข้อมูลโครงการ เมื่อผู้ใช้งานเลือกจะเข้าสู่หน้าแสดงรูปภาพแผนผังโครงการ โดยระบบจะแสดงข้อมูลแบบบ้าน ที่ดิน ราคา และสถานะการขาย เป็นรายแปลง เมื่อผู้ใช้งานเลือกแปลงที่ต้องการ จะแสดงหน้าต่าง Popup เพื่อให้ดาวน์โหลดไฟล์โมเดลแบบบ้าน สำหรับแสดงด้วยเทคนิคความจริงเสริม โดยถ้าต้องการแสดงผ่าน Smart Phone ต้องสแกนผ่าน QR Code หรือถ้าใช้คอมพิวเตอร์ตั้งโต๊ะสามารถกดดาวน์โหลดและแสดงผลได้ทันทีผ่านโปรแกรม AR Player จากนั้นนำแผ่นสัญลักษณ์มาส่งที่หน้ากล้อง โดยโมเดลบ้านจะแสดงตรงตำแหน่งแผ่นสัญลักษณ์ที่ผู้ใช้งานถือไว้ ซึ่งผู้ใช้งานจะสามารถมีปฏิสัมพันธ์กับโมเดลบ้านได้ในรูปแบบต่าง ๆ เช่น ย่อขยายขนาด หมุนโมเดล เป็นต้น



ภาพที่ 1: แผนผังการทำงานของผู้ใช้งานระบบ

3.3 การพัฒนาระบบ

เป็นกระบวนการในการพัฒนาระบบให้ได้ตามข้อกำหนดที่ได้ออกแบบไว้ เพื่อให้ออกมาเป็นระบบที่สามารถใช้งานได้ ซึ่งการพัฒนาระบบช่วยนำเสนอข้อมูลแบบบ้านจัดสรรด้วยเทคนิคความจริงเสริมนี้ ผู้วิจัยได้ใช้โปรแกรมภาษา C# (ASP.NET) ที่อยู่ในชุดของโปรแกรม Microsoft Visual Studio Express 2012 ในการสร้างเว็บเพจ และหน้าที่ติดต่อกับผู้ใช้งาน ร่วมกับ JavaScript เพื่อให้ระบบสามารถทำงานได้อย่างมีประสิทธิภาพอย่างสูงสุด โปรแกรมจัดการฐานข้อมูลได้เลือกใช้ Microsoft SQL Server 2008 Express Edition ซึ่งจะติดต่อโดยผ่าน ADO.NET ส่วนการตกแต่งภาพและสร้างโมเดลแบบบ้าน 3 มิติได้เลือกใช้โปรแกรม Adobe Photoshop CS5, Google SketchUp 8 และการแสดงผลด้วยเทคนิคความจริงเสริมเลือกใช้ AR-media™ ในการพัฒนา ทั้งนี้ผู้วิจัยได้แบ่งการพัฒนาส่วนหลัก ๆ ออกเป็น 3 ส่วน ได้แก่ การจัดการข้อมูลโครงการและสินค้า การแสดงผลข้อมูลโครงการและสินค้า การทำแผ่นสัญลักษณ์ (Marker) และการทำโมเดลแบบบ้าน 3 มิติ ที่ใช้ในการแสดงผลข้อมูลด้วยเทคนิคความจริงเสริม

3.4 การทดสอบระบบ

เป็นกระบวนการในการทดสอบระบบที่พัฒนาขึ้น เพื่อให้ทำงานได้อย่างถูกต้องตรงความต้องการ โดยนำระบบมาทดสอบระดับหน่วยย่อย และการใช้งานระบบภายใต้สภาพแวดล้อมการใช้งานจริง แล้วนำปัญหาจากการทดสอบมาปรับปรุงแก้ไขเพื่อให้การทำงานสมบูรณ์มากที่สุด

3.5 การประเมินผลระบบ

เมื่อทำการปรับปรุงแก้ไขเสร็จเรียบร้อยแล้วจึงนำระบบที่พัฒนาขึ้นไปทดสอบกับกลุ่มเป้าหมาย เพื่อประเมินความพึงพอใจของระบบจากผู้ใช้งานทั่วไป 15 คน และผู้เชี่ยวชาญ 5 คน รวมทั้งหมด 20 คน โดยทางผู้วิจัยได้ออกแบบและสร้างแบบสอบถามเพื่อใช้ในการประเมิน นอกจากนี้ยังได้ประเมินประสิทธิภาพการทำงานของระบบควบคู่กัน จากนั้นจึงนำข้อมูลที่ได้ออกมาวิเคราะห์ผลลัพท์ต่อไป

4. ผลการดำเนินงาน

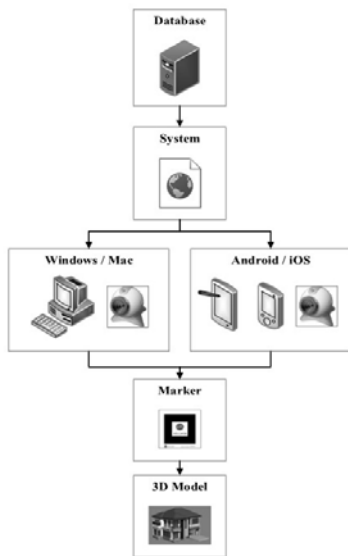
ผลการดำเนินงานของงานวิจัยครั้งนี้ได้ทำการทดสอบประสิทธิภาพของระบบที่พัฒนาขึ้น และประเมินความพึงพอใจ

ของระบบ ซึ่งผู้วิจัยได้ประเมินจากผู้ใช้งานทั่วไป และผู้เชี่ยวชาญ โดยมีรายละเอียดดังต่อไปนี้

4.1 รูปแบบและการทำงานของระบบ

ระบบแสดงข้อมูลแบบบ้านจัดสรรด้วยเทคนิคความจริงเสริมเริ่มจากการพัฒนาแบบบ้าน 3 มิติ 5 แบบบ้าน และปรับรูปแบบไฟล์ให้ใช้งานได้กับ AR Player

จากนั้นผู้ใช้งานเริ่มต้นการใช้งานจากหน้าหลักของโครงการ ซึ่งสามารถเลือกสถานะที่ต้องการแสดงผลได้ 3 สถานะ ได้แก่ สถานะขาย แบบบ้าน และราคา โดยระบบเป็นลักษณะเว็บแอปพลิเคชัน เพื่อให้การใช้งานง่ายต่อผู้ใช้งานมากที่สุด ระบบสามารถเข้าถึงได้จากทุกที่



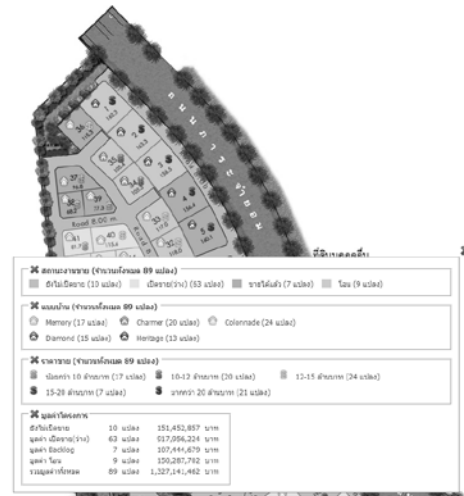
ภาพที่ 2: ภาพรวมการทำงานของระบบ



ภาพที่ 3: ภาพหน้าหลักโครงการของระบบ

เมื่อผู้ใช้งานเลือกสถานะที่ต้องการเรียบร้อยแล้ว ระบบจะแสดงข้อมูลแผนผังโครงการพร้อมสถานะตามที่เราเลือกไว้ ดังภาพที่ 4 อีกทั้งผู้ใช้งานสามารถเปิด-ปิดสถานะข้อมูลต่าง ๆ

ได้อย่างอิสระ เพื่อให้วิเคราะห์ข้อมูลที่ต้องการได้อย่างมีประสิทธิภาพ



ภาพที่ 4: ภาพแสดงสถานะของโครงการ

จากนั้นเมื่อผู้ใช้งานเลือกแปลงที่ต้องการ ระบบจะแสดงหน้าต่างป๊อปอัพข้อมูลต่าง ๆ ได้แก่ ขนาดที่ดิน ชื่อแบบบ้าน พื้นที่ใช้สอย ราคา รายละเอียดแบบบ้าน พร้อมทั้งวิธีการแสดงผลแบบบ้านด้วยเทคนิคความจริงเสริมเป็นลำดับขั้นตอน โดยเริ่มจากดาวน์โหลด AR Player สำหรับอุปกรณ์ที่ต้องการใช้แสดงผลขั้นถัดไปต้องพิมพ์แผ่นสัญลักษณ์หรือเปิดด้วย Tablet จากนั้นถ้าใช้งานด้วย iOS / Android ให้สแกน QR Code ด้วย AR Player หรือถ้าเป็น Windows / Mac ให้กดดาวน์โหลดไฟล์ AR เพื่อแสดงผลแบบบ้านตามวิธีที่ต้องการ และเริ่มต้นใช้งานได้ที่



ภาพที่ 5: หน้าต่างแสดงข้อมูลสินค้า



ภาพที่ 6: หน้าต่างแสดงผลแบบบ้านด้วยเทคนิคความจริงเสริม

4.2 ผลการทดสอบประสิทธิภาพ

หลังจากพัฒนาระบบเสร็จเรียบร้อยแล้วได้ทำการทดสอบประสิทธิภาพ โดยผลการทดสอบมีรายละเอียดดังต่อไปนี้

ตารางที่ 1: ผลการทดสอบประสิทธิภาพ

หัวข้อ	จำนวนครั้ง	ผ่าน	ไม่ผ่าน
1. การแสดงผลแบบบ้านด้วยเทคนิคความจริงเสริม ในระยะ 50 ซม.	100	98	2
2. การแสดงผลแบบบ้านด้วยเทคนิคความจริงเสริม ในระยะ 100 ซม.	100	95	5
3. ข้อมูลแบบบ้านรายแปลงในหน้าต่างป๊อปอัพ	87	87	0
รวม	287	280	7

จากตารางที่ 1 ผลการทดสอบประสิทธิภาพแต่ละหัวข้อการทดสอบพบว่า การแสดงผลแบบบ้านด้วยเทคนิคความจริงเสริมในระยะ 50 ซม. แสดงผลได้ถูกต้อง 98 ครั้ง จาก 100 ครั้ง ส่วนการแสดงผลแบบบ้านด้วยเทคนิคความจริงเสริม ในระยะ 100 ซม. แสดงผลได้ถูกต้อง 95 ครั้ง จาก 100 ครั้ง และข้อมูลแบบบ้านรายแปลงในหน้าต่างป๊อปอัพมีความถูกต้องทั้ง 87 แปลง เมื่อรวมคะแนนทั้งหมดจะได้ 280 จาก 287 คิดเป็นร้อยละ 97.56 ดังนั้นระบบที่พัฒนาขึ้นมีประสิทธิภาพอยู่ในระดับดีมาก

4.3 ผลการประเมินความพึงพอใจ

ผลการประเมินความพึงพอใจระบบของผู้ใช้งานทั่วไป 15 คน และผู้เชี่ยวชาญ 5 คน รวมทั้งหมด 20 คน มีรายละเอียดดังต่อไปนี้

ตารางที่ 2: ผลการประเมินความพึงพอใจ

หัวข้อ	Mean	S.D.
1. ระยะเวลาในการแสดงผล	4.80	0.44
2. การตอบสนองของแผ่นสัญลักษณ์	4.40	0.54
3. ความถูกต้องของข้อมูล	4.80	0.44
4. เนื้อหาชัดเจนเข้าใจง่าย	4.80	0.44
5. ขนาดของตัวอักษรที่ใช้งาน	4.40	0.54
6. การใช้รูปภาพในภาพรวม	4.40	0.54
7. การใช้สีในภาพรวม	4.40	0.54
8. ความดึงดูดและน่าสนใจ	4.80	0.44
9. ความยากง่ายในการใช้งาน	4.20	0.83
10. คำแนะนำวิธีต่าง ๆ	4.20	0.83
รวม	4.52	0.56

จากตารางที่ 2 ผลการประเมินความพึงพอใจ เมื่อใช้ค่าเฉลี่ย (Mean) และค่าส่วนเบี่ยงเบนมาตรฐาน (S.D.) จะพบว่าค่าเฉลี่ยได้อยู่ที่ 4.52 และมีค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.56 ดังนั้นระบบที่พัฒนาขึ้นมีค่าเฉลี่ยความพึงพอใจที่มีต่อระบบมีค่าอยู่ในระดับดีมาก

5. สรุปผล อภิปรายผล และข้อเสนอแนะ

งานวิจัยในครั้งนี้มีวัตถุประสงค์เพื่อพัฒนาระบบนำเสนอข้อมูลแบบบ้านจัดสรรโดยนำเทคนิคความจริงเสริมมาการประยุกต์ใช้งาน โดยเริ่มต้นจากการศึกษาข้อมูลของเทคนิคความจริงเสริม ขั้นตอนการพัฒนาระบบ รวมไปถึงเทคโนโลยีที่เกี่ยวข้อง จากนั้นวิเคราะห์ข้อมูลที่จะใช้ กำหนดขอบเขตออกแบบ และพัฒนาเป็นเว็บแอปพลิเคชัน โดยใช้โปรแกรมภาษา C# (ASP.NET) ที่อยู่ในชุดของโปรแกรม Microsoft Visual Studio Express 2012 ในการสร้างเว็บเพจ และหน้าข้อที่ติดต่อกับผู้ใช้งานร่วมกับ JavaScript ส่วนโปรแกรมจัดการฐานข้อมูลเลือกใช้ Microsoft SQL Server 2008 Express Edition ซึ่งจะติดต่อโดยผ่าน ADO.NET การตกแต่งภาพและสร้างโมเดลแบบบ้าน 3 มิติเลือกใช้โปรแกรม Adobe Photoshop CS5, Google SketchUp 8 และการแสดงผลด้วยเทคนิคความจริงเสริมเลือกใช้ AR-media™ ซึ่งข้อดีของการพัฒนาเป็นเว็บแอปพลิเคชันจะช่วยให้ผู้ใช้งานสามารถเข้าใช้งานได้หลากหลายอุปกรณ์ รวมไปถึงรูปแบบ

เทคนิคความจริงเสริมที่เลือกใช้ก็สามารถรองรับหลายอุปกรณ์ โดยในการใช้งานเทคนิคความจริงเสริมต้องใช้ควบคู่กับแผ่นสัญลักษณ์และกล้องเว็บแคม

จากนั้นเมื่อพัฒนาระบบเสร็จเรียบร้อยแล้ว จึงนำระบบไปทดสอบประสิทธิภาพและประเมินความพึงพอใจกับผู้ใช้งานทั่วไป 15 คน และผู้เชี่ยวชาญ 5 คน รวมทั้งหมด 20 คน โดยใช้ค่าเฉลี่ยและค่าส่วนเบี่ยงเบนมาตรฐาน โดยมีรายละเอียดดังต่อไปนี้

5.1 สรุปผล

หลังจากพัฒนาระบบเสร็จเรียบร้อยแล้ว สามารถนำผลการทดสอบประสิทธิภาพและประเมินความพึงพอใจมาสรุป โดยประสิทธิภาพมาจากการทดสอบการแสดงผลแบบบ้านด้วยเทคนิคความจริงเสริมในระยะ 50 ซม. และ 100 ซม. รวมไปถึงการแสดงผลข้อมูลแบบบ้านเป็นรายแปลงในหน้าต่างป๊อปอัพของระบบ ซึ่งได้ผลคะแนนรวมคิดเป็นร้อยละ 97.56 จัดอยู่ในระดับดีมาก ส่วนความพึงพอใจ ได้พิจารณาจากหัวข้อต่าง ๆ ได้แก่ ระยะเวลาในการแสดงผล การตอบสนองของแผ่นสัญลักษณ์ ความถูกต้องของข้อมูล เนื้อหาชัดเจนเข้าใจง่าย ขนาดของตัวอักษรที่ใช้ การใช้งานรูปภาพในภาพรวม การใช้สีในภาพรวม ความดึงดูดและน่าสนใจ ความง่ายในการใช้งาน และคำแนะนำวิธีต่าง ๆ โดยได้ค่าเฉลี่ย 4.52 และค่าส่วนเบี่ยงเบนมาตรฐาน 0.56 จึงสรุปได้ว่าความพึงพอใจของผู้ใช้งานอยู่ในระดับดีมาก

5.2 อภิปรายผล

จากการพัฒนาระบบแสดงผลข้อมูลแบบบ้านจัดสรรด้วยเทคนิคความจริงเสริม แล้วนำไปทดสอบประสิทธิภาพและประเมินความพึงพอใจผู้ใช้งาน โดยผลที่ได้ออกมาเป็นคะแนนอยู่ในระดับดีมากทั้งประสิทธิภาพและความพึงพอใจ ซึ่งจากการนำไปให้ผู้ใช้งานทดสอบพบว่าผู้ใช้งานมีความสนใจ และเกิดความรู้สึกแปลกใหม่ในการนำเสนอข้อมูล ในส่วนของปัญหาของการใช้งานระบบที่พบคือ ในบางครั้งอุปกรณ์ที่ใช้งานมีคุณภาพของกล้องไม่ดีพอทำให้ระบบตรวจจับแผ่นสัญลักษณ์ไม่ชัดเจน รวมไปถึงปริมาณแสงที่มากเกินไปก็จะส่งผลกระทบต่อตรวจจับเช่นกัน แต่ในภาพรวมแล้วผู้ใช้งานก็ได้ให้คะแนนความพึงพอใจอยู่ในระดับดีมาก รวมไปถึงคะแนนด้านประสิทธิภาพก็ได้คะแนนที่สูงมากถึง 97.56% ดังนั้นระบบที่พัฒนาขึ้นจึงสามารถนำไปใช้งานได้จริง ซึ่งจะช่วยให้เกิดความแปลกใหม่น่าสนใจ

สามารถเห็นรูปแบบบ้านเหมือนจริงอันจะช่วยให้เกิดการตัดสินใจในด้านต่าง ๆ ได้อย่างแม่นยำและมีประสิทธิภาพต่อไป

5.3 ข้อเสนอแนะ

จากการวิจัยครั้งนี้ทำให้ทราบว่าหากต้องการพัฒนาระบบต่อไปในอนาคตนั้น ส่วนที่สำคัญในการพัฒนาต่อคือรูปแบบของข้อมูลที่จะนำมาแสดงผล ควรมีความน่าสนใจมากขึ้นไปอีก เช่นสามารถกดเปิดหลังคาบ้านเพื่อดูการจัดแบ่งสัดส่วนต่าง ๆ ของบ้าน การจัดวางเฟอร์นิเจอร์ภายในบ้าน การแสดงข้อมูลขนาดต่าง ๆ ของพื้นที่บ้าน เป็นต้น ส่วนข้อมูลบนเว็บแอปพลิเคชันต้องมีความครบถ้วนหลากหลายมากขึ้น และยังสามารถนำไปประยุกต์ใช้กับอุปกรณ์แบบสวมหัวเพื่อให้สามารถใช้งานได้ง่ายและมีประสิทธิภาพอย่างสูงสุด นอกจากนี้ยังสามารถนำไปต่อยอดศึกษาผลกระทบของการนำมาใช้ในธุรกิจอีกด้วย

เอกสารอ้างอิง

- [1] พนิดา ดันศิริ, “โลกเสมือนผสมผสานโลกจริง”, มหาวิทยาลัยกรุงเทพ, 2550.
- [2] PSFK, “App Shows Road Accidents In Augmented Reality”, 2011. Available online at : <http://www.psfk.com/2011/08/russianapp-shows-road-accidents-in-augmented-reality-video.html>
- [3] Jupiter Research, “Mobile Augmented Reality Users to Approach 200 Million Globally by 2018”, 2013. Available online at : <http://www.juniperresearch.com/viewpressrelease.php?pr=410>
- [4] สุพรรณพงศ์ วงษ์ศรีเพ็ง, “การประยุกต์ใช้เทคนิคความจริงเสริมเพื่อใช้ในการสอนเรื่องพญชุนะภาษาไทย”, คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2554.
- [5] มงคล หล้าดวงดี, “ระบบสารสนเทศด้านการท่องเที่ยวด้วยความเป็นจริงเสริมบนโทรศัพท์เคลื่อนที่”, สาขาวิศวกรรมซอฟต์แวร์ มหาวิทยาลัยเชียงใหม่, 2553.
- [6] ณัฐพล ปฐมอริย์, “ระบบความจริงเสริมสำหรับการถ่ายทอดทักษะการประกอบชิ้นงาน”, คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี, 2547.
- [7] พีรศิลป์ สันธนะพันธ์, “ระบบความจริงเสริมบนสภาพแวดล้อมที่เน้นการสัมผัสสำหรับการเรียนรู้”, สถาบันวิทยาการหุ่นยนต์ภาคสนาม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี, 2551.
- [8] Henderson, S. and Feiner S, “Augmented Reality for Maintenance and Repair (ARMAR)”, Department of Computer Science, Columbia University, 2007.

การทำภาพพืชมัลลายน้ำดิจิทัลโดยใช้โดเมนเวฟเลตและโดเมนเวลา

Image Watermarking using Joined Wavelet and Time Domain

อักรพล ผดุงดิษฐ์ (Akarapol Padungdit)¹ สิปปวิชญ์ สุขกรี (Sippawit Suggree)² และ ปานวิทย์ ฐะวะนุติ (Panwit Tuwanut)³
คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

¹itschmp@gmail.com, ²ipunnnnnn@gmail.com, ³panwit@it.kmitl.ac.th

บทคัดย่อ

บทความฉบับนี้ กล่าวถึงการทำภาพพืชมัลลายน้ำดิจิทัลสำหรับภาพสี โดยใช้เทคนิคการฝังลายน้ำในโดเมนเวลาและโดเมนความถี่ ด้วยการแปลงเวฟเลตในเวลาเดียวกัน กระบวนการที่นำเสนอมีข้อได้เปรียบกว่ามากกว่ากระบวนการฝังลายน้ำลงในโดเมนเวลาหรือโดเมนความถี่ด้วยการแปลงเวฟเลตเพียงโดเมนเดียว ในแบบเดิมคือลายน้ำที่ถูกฝังลงในภาพต้นฉบับจะมีความทนทานต่อการโจมตีทั่วไปในหลายๆด้าน ทั้งการโจมตีเชิงเรขาคณิต เช่น การย่อขนาด การตัดขอบภาพออก หรือการโจมตีเชิงสัญญาณ เช่น การบีบอัดภาพ การทดสอบได้แสดงให้เห็นว่ากระบวนการที่ได้นำเสนอนี้ให้ลายน้ำที่กู้คืนมาหลังจากผ่านการโจมตีในหลายๆรูปแบบดีกว่าการฝังลายน้ำลงในโดเมนเวลาหรือโดเมนความถี่ด้วยการแปลงเวฟเลตเพียงโดเมนเดียว

คำสำคัญ: ลายน้ำดิจิทัล, การฝังข้อมูล, ภาพสี, เวฟเลต

Abstract

This paper presents Image Watermarking Using Joined Wavelet and Time Domain. This method proposed has advantages above a single domain watermarking which are time domain or frequency domain. The advantages are that the output watermarked image is more robust and the result watermark after attacked from both geometric attacks and frequency attacks is better than the single watermarking method.

Keywords: Digital Watermarking, Data Embedded, Color Image, Wavelet Transform

1. บทนำ

ในปัจจุบันการติดต่อสื่อสารสามารถทำได้อย่างรวดเร็ว ทำให้ปัญหาทางการละเมิดลิขสิทธิ์ทำได้ง่าย ดังนั้นวิธีการติดตามและยืนยันความเป็นเจ้าของสิทธิ์จึงเป็นอีกวิธีหนึ่งที่ช่วยให้เจ้าของสิทธิ์นั้นสามารถมั่นใจได้ว่าผู้อื่นไม่สามารถหาผลประโยชน์จากผลงานของตนเองได้ การทำภาพพืชมัลลายน้ำดิจิทัลจึงเข้ามามีบทบาทที่ช่วยในการยืนยันสิทธิ์ได้เป็นอย่างดี

ภาพพืชมัลลายน้ำดิจิทัล เป็นเทคนิคในการตรวจสอบความเป็นเจ้าของในผลงาน โดยการซ่อนข้อมูลที่ใช้ในการยืนยันสิทธิ์ความเป็นเจ้าของกับผลงานที่สร้างสรรค์ขึ้น โดยสามารถใช้ได้กับข้อมูลหลายๆชนิด เช่น รูปภาพ, เสียง, วิดีโอ ได้อย่างมีประสิทธิภาพ ซึ่งภาพพืชมัลลายน้ำดิจิทัลที่ดีนั้น จะต้องไม่ทำให้คุณภาพของรูปภavnั้นลดต่ำลงจนเกินไปและเมื่อมีการคัดลอกข้อมูลเกิดขึ้น ลายน้ำดังกล่าวต้องติดไปกับข้อมูลที่มีการคัดลอกไปด้วย

เทคนิคในการทำภาพพืชมัลลายน้ำดิจิทัลอาจถูกแยกประเภทออกได้ตามโดเมนที่กระทำการฝังลายน้ำดิจิทัล เช่น โดเมนเวลาหรือโดเมนความถี่ สัญญาณลายน้ำต่างๆอาจสูญหายได้เมื่อผ่านการโจมตีต่างๆ โดเมนเวลาจะสูญหายเมื่อผ่านการโจมตีทางสัญญาณ เช่น การกรองสัญญาณในแบบต่างๆ แต่จะทนทานต่อการโจมตีเชิงเรขาคณิต เช่น การย่อภาพ การตัดภาพ เป็นต้น ในทางกลับกัน การฝังลายน้ำลงในโดเมนความถี่จะทำให้สัญญาณลายน้ำทนทานต่อการโจมตีเชิงสัญญาณ เช่น การกรองความถี่ต่ำ การกรองความถี่สูง เป็นต้น จะเห็นได้ว่าการฝังลายน้ำลงในโดเมนที่ต่างกัน ย่อมทนทานต่อการโจมตีในรูปแบบกัน ทั้งนี้หากภาพพืชมัลลายน้ำดิจิทัลถูกโจมตีในหลายๆแบบ ย่อมอาจทำให้ลายน้ำสูญเสียได้

ดังนั้น งานวิจัยฉบับนี้จึงทำการศึกษาและพัฒนาการฝังลายน้ำทั้งโดเมนเวลาและโดเมนความถี่ด้วยการแปลงเวฟเลตในเวลาเดียวกัน เพื่อให้ภาพพิมพ์ลายน้ำดิจิทัลทนทานต่อการโจมตีทั้งสองประเภท คือ การโจมตีเชิงสัญญาณ และการโจมตีเชิงเรขาคณิตในเวลาเดียวกัน ซึ่งการโจมตีเหล่านี้มีโอกาสเกิดขึ้นได้ในการสื่อสารทางอินเทอร์เน็ตในปัจจุบัน โดยใช้ภาพต้นฉบับและภาพลายน้ำอันเดียวกันในทุกวิธีการ เปรียบเทียบความผิดเพี้ยนของภาพลายน้ำที่กู้คืนมาหลังผ่านการโจมตีกับภาพลายน้ำต้นฉบับ

2. วรรณกรรมที่เกี่ยวข้อง

โดยทั่วไปแล้วงานวิจัยส่วนใหญ่จะมุ่งเน้นในการหาตำแหน่งที่จะใส่สัญญาณลายน้ำดิจิทัล ซึ่งจะทำให้ลายน้ำที่ถูกฝังลงในภาพต้นฉบับมีความทนทานต่อการโจมตีพื้นฐาน ซึ่งมักจะฝังในโดเมนเดียว

การทำภาพพิมพ์ลายน้ำดิจิทัลที่อ้างอิงจากเทคนิคการกระจายแถบความถี่แบบอันดับตรง (Direct Sequence Spread Spectrum) ได้ถูกเสนอโดย M. George และคณะ [1] ด้วยวิธีการนี้สัญญาณลายน้ำจะถูกขยายออกด้วยค่าคงที่หนึ่งเรียกว่า ชิปรต (Chip Rate) ทำให้ได้สัญญาณลายน้ำที่ถูกกระจายแถบความถี่ จากนั้นนำข้อมูลไปมอดูเลตกับสัญญาณรบกวนแบบสุ่มเทียม (Pseudo Random Noise) แล้วปรับค่าความแรงสัญญาณด้วยค่าแอมพลิจูดแพกเตอร์ที่ทำให้ลายน้ำมีความคงทนและปลอดภัยมากขึ้น

ต่อมา T. Amornraksa และคณะ [2] ได้นำเสนอวิธีการในการทำภาพพิมพ์ลายน้ำบนโดเมนเวลา ซึ่งเทคนิคการกระจายความถี่แบบกระโดด (Frequency Hopping Spread Spectrum) และการกระจายแถบความถี่แบบอันดับตรงถูกนำมาใช้ร่วมกัน เนื่องจากคุณสมบัติเฉพาะตัวของเทคนิคนี้ทำให้ได้ลายน้ำที่ยากต่อการหาตำแหน่งและตรวจจับ

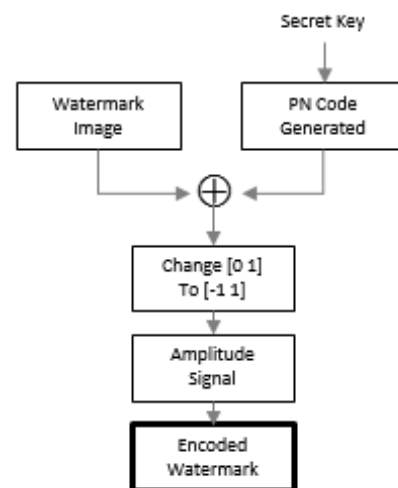
สำหรับการทำภาพพิมพ์ลายน้ำในโดเมนความถี่ด้วยการแปลงเวฟเลตร่วมกับเทคนิคการกระจายแถบความถี่อันดับตรง ถูกนำเสนอโดย Yu Pin Wang และคณะ [3] ซึ่งได้นำเสนอวิธีการทำภาพพิมพ์ลายน้ำบนโดเมนความถี่ ด้วยการแปลงเวฟเลตแบบไม่ต่อเนื่อง (Discrete Wavelet Transform: DWT) โดยสัญญาณลายน้ำจะถูกนำไปแทรกลงในช่วงความถี่ต่ำของค่า

สัมประสิทธิ์ที่ได้จากการแปลงเวฟเลตแบบไม่ต่อเนื่องของภาพต้นฉบับ ซึ่งผลการทดลองแสดงให้เห็นว่าภาพที่ได้ลายน้ำด้วยวิธีการข้างต้น เมื่อผ่านกระบวนการบีบอัดข้อมูลด้วยอัตราบีบอัดสูง ๆ จะได้สัญญาณลายน้ำผลลัพธ์ที่ทนทานกว่าการแปลงโคไซน์แบบไม่ต่อเนื่อง

3. วิธีการดำเนินการวิจัย

วิธีการที่นำเสนอในที่นี้จะกระทำการฝังลายน้ำในสัญญาณแสงของภาพจำนวน 2 ช่องด้วยกัน คือ สัญญาณภาพสีแดงบนโดเมนความถี่ และสัญญาณภาพสีน้ำเงินบนโดเมนเวลา โดยภาพต้นฉบับจะใช้ภาพสีขนาด 512x512 พิกเซล และลายน้ำเป็นภาพขาวดำขนาด 256x256 พิกเซล

ในการเตรียมลายน้ำ จะเริ่มจากการขยายบิตสัญญาณลายน้ำโดยวิธีการกระจายแถบความถี่แบบอันดับตรง (Direct Sequence Spread Spectrum) โดยการสร้างเลขสุ่มเทียม (Pseudo Random Noise Generator) ขนาด 4x4 บิตขึ้นมา จากนั้นทำให้ได้ขนาดเท่ากับลายน้ำ คือ 256x256 บิต แล้วทำการ XOR กับลายน้ำเพื่อให้ได้ลายน้ำที่ผ่านการเข้ารหัสแล้วนั่นเอง แผนภาพการเข้ารหัสลายน้ำแสดงดังภาพที่ 1



ภาพที่ 1: แผนภาพการเข้ารหัสลายน้ำ

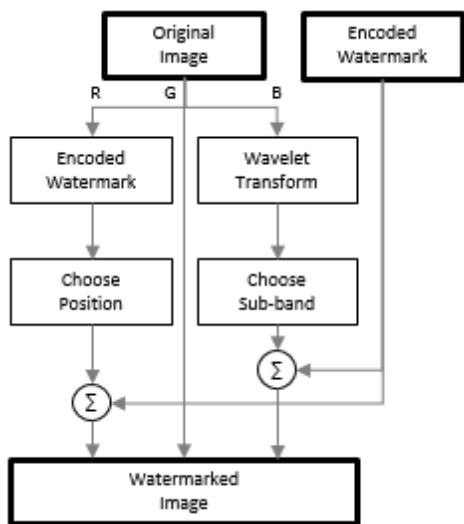
เมื่อได้ลายน้ำที่เข้ารหัสเรียบร้อยแล้ว ก็จะทำการเพิ่มความแรงของลายน้ำด้วยการขยายแอมพลิจูดของสัญญาณ เพื่อให้ลาย

น้ำมีความแรงมากขึ้นเมื่อส่งลงในภาพต้นฉบับ การลายน้ำจะส่งลงไปในช่วงสัญญาณสีสองช่องด้วยกันดังที่ได้กล่าวไว้ข้างต้น

ช่องสัญญาณสีแดงทำการฝังลายน้ำบนโดเมนความถี่ด้วยการแปลงเวฟเลต เลือกเฉพาะแบนด์ย่อยความถี่ต่ำที่สุดทั้งแกนนอนและแกนตั้ง (LL Sub-band) เพราะทนทานต่อการประมวลผลแบบบีบอัดข้อมูลภาพได้ดีที่สุด จากนั้นนำภาพลายน้ำที่เข้ารหัสและปรับขยายแอมพลิจูดแล้วบวกกับแบนด์ย่อย LL ตรง ๆ แล้วทำการแปลงกลับเวฟเลต ก็จะได้ภาพฝังลายน้ำออกมา

ช่องสัญญาณสีน้ำเงินทำการฝังลายน้ำบนโดเมนเวลาโดยการเลือกจุดที่ฝังนั้นจะใช้เทคนิคการกระจายแถบความถี่แบบความถี่กระโดด (Frequency Hopping Spread Spectrum Technique) ด้วยการสุ่มพิกเซลทุกจุดของลายน้ำแล้วทำการสเกลความแรงของลายน้ำให้ลดลง เพื่อให้ยากต่อการตรวจจับด้วยสายตามนุษย์แล้วนำไปบวกตรง ๆ กับภาพต้นฉบับ ก็จะได้ภาพฝังลายน้ำอีกภาพหนึ่ง

เมื่อทำการฝังทั้งสองโดเมน สองช่องสัญญาณ จะได้ภาพฝังลายน้ำที่มีสัญญาณรบกวนจากการฝังลายน้ำออกมา แผนภาพการแสดงผลการฝังลายน้ำทั้งสองโดเมนแสดงดังภาพที่ 2



ภาพที่ 2: แผนภาพการฝังลายน้ำทั้งโดเมนความถี่และเวลา

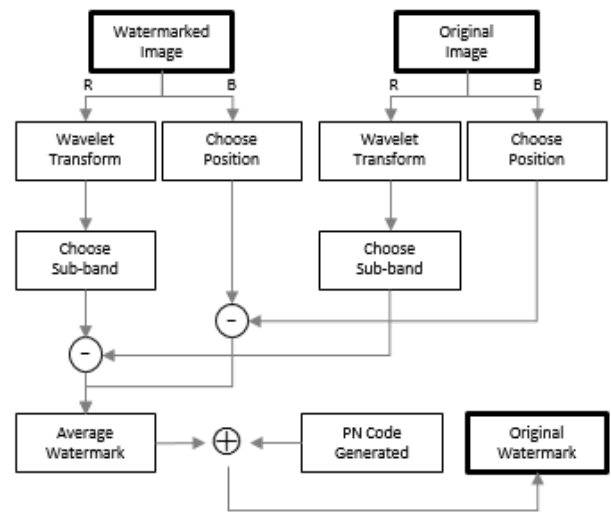
หลังจากได้ผลลัพธ์การฝังลายน้ำแล้วก็จะทำการถอดลายน้ำออก เพื่อทำการเปรียบเทียบความแตกต่างระหว่างภาพลายน้ำ

ต้นฉบับ กับลายน้ำที่ถอดออกได้ เพื่อวิเคราะห์ประสิทธิภาพของการฝังลายน้ำด้วยวิธีการที่นำเสนอ

ช่องสัญญาณสีแดงที่ทำการฝังบนโดเมนความถี่ ต้องทำการแปลงความถี่เวฟเลตทั้งภาพต้นฉบับและภาพที่ถูกฝังลายน้ำแล้วเสียก่อน จากนั้นจึงนำแบนด์ย่อยความถี่ต่ำของทั้งสองภาพมาลบกัน โดยให้ภาพฝังลายน้ำเป็นตัวตั้งและลบด้วยภาพต้นฉบับ จะได้ลายน้ำที่ถูกฝังอยู่ออกมา

ช่องสัญญาณสีน้ำเงินที่ทำการฝังบนโดเมนเวลา จะนำภาพต้นฉบับกับภาพฝังลายน้ำมาเลือกเฉพาะจุดพิกเซลที่ฝังลายน้ำลงไป ลบกันโดยตรงจะได้ภาพลายน้ำขนาด 512x512 พิกเซลที่กระจายกันอยู่ออกมา จากนั้นนำไปจัดเรียงให้ได้กลับเป็นแบบลายน้ำต้นฉบับเดิม

เมื่อได้ลายน้ำของทั้งสองโดเมนแล้วก็จะทำการเฉลี่ยเพราะลายน้ำอาจเสียหายจากการโจมตีได้ แล้วจึงนำไป XOR กับเลขสุ่มเทียมที่สร้างขึ้นเพื่อทำให้ได้ลายน้ำต้นฉบับจริง ๆ ออกมา แผนภาพการถอดลายน้ำของทั้งสองโดเมนแสดงดังภาพที่ 3



ภาพที่ 3: แผนภาพการถอดลายน้ำทั้งโดเมนความถี่และเวลา

4. ผลการดำเนินงาน

ในการทดสอบลายน้ำ จำเป็นต้องมีการขยายความของตัวแปรที่ใช้ในการวัดผลบางตัว เพื่อความเข้าใจในผลการทดลองมากขึ้น

- Peak Signal to Noise Ratio (PSNR) – ในการหาค่าความผิดพลาดในเชิงตัวเลขของการประมวลผลภาพนิยมใช้ค่าความผิดพลาดกำลังสองเฉลี่ย (Mean Square Error - MSE) ถ้ามีค่าน้อยแสดงว่าภาพที่ถูกส่งลายน้ำมีความผิดเพี้ยนจากต้นฉบับน้อย สามารถหาค่าได้ดังสมการที่ 1

$$MSE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N [W(i,j) - O(i,j)]^2 \quad (1)$$

โดย MSE คือ ค่าความผิดพลาดกำลังสองเฉลี่ย
 M คือ จำนวนพิกเซลตามความกว้างภาพ
 N คือ จำนวนพิกเซลตามความยาวภาพ
 W(i,j) คือ ค่าพิกเซลที่ตำแหน่ง (i,j) ของภาพที่ถูกส่งลายน้ำ
 O(i,j) คือ ค่าพิกเซลที่ตำแหน่ง (i,j) ของภาพต้นฉบับ

เพื่อให้ได้เครื่องมือวัดเชิงตัวเลขที่ชัดเจนและชัดเจนยิ่งขึ้น เราเลือกใช้อัตราส่วนของสัญญาณต่อสัญญาณรบกวนสูงสุด (PSNR: Peak Signal to Noise Ratio) ในการวัดสัญญาณรบกวนในภาพส่งลายน้ำ ซึ่งเป็นวิธีที่นิยมเป็นอย่างแพร่หลาย ค่า PSNR สามารถหาได้จากสมการที่ 2

$$PSNR = 10 \log_{10} \frac{Peak^2}{MSE} (dB) \quad (2)$$

โดย PSNR คือ อัตราส่วนของสัญญาณต่อสัญญาณรบกวนสูงสุด
 MSE คือ ค่าความผิดพลาดกำลังสองเฉลี่ย
 Peak คือ ค่าสูงสุดของข้อมูลซึ่งในที่นี้คือ 255 เพราะค่าในแต่ละพิกเซลของภาพ RGB คือ 0-255

- Normalize Correlation (NC) – นอกจากการคำนวณหาค่า PSNR ในการวัดค่าความผิดเพี้ยนของภาพที่ถูกส่งลายน้ำแล้ว การเปรียบเทียบลายน้ำที่ตรวจหาได้กับลายน้ำต้นฉบับยังสามารถทำการวัดค่า

เปรียบเทียบได้ โดยหลังจากที่ถูกลายน้ำกลับคืนมาแล้วสามารถหาค่า NC ได้จากสมการที่ 3

$$NC = \frac{\sum_{i=1}^m \sum_{j=1}^n w(i,j)w'(i,j)}{\sum_{i=1}^m \sum_{j=1}^n w(i,j)^2} \times 100(\%) \quad (3)$$

โดย NC คือ ค่าความสัมพันธ์ของลายน้ำที่ตรวจพบกับลายน้ำต้นฉบับ
 m คือ จำนวนพิกเซลตามความกว้างของภาพลายน้ำ
 n คือ จำนวนพิกเซลตามความยาวของภาพลายน้ำ
 w'(i,j) คือ ค่าพิกเซลที่ตำแหน่ง (i,j) ของภาพลายน้ำที่ตรวจพบ
 w''(i,j) คือ ค่าพิกเซลที่ตำแหน่ง (i,j) ของภาพลายน้ำต้นฉบับ

ค่า NC ที่ได้แสดงถึงความเหมือนของภาพลายน้ำต้นฉบับกับลายน้ำที่ตรวจพบ โดยค่าที่ได้หากมีค่าใกล้เคียง 100% แสดงว่าลายน้ำที่ตรวจพบมีความเหมือนกับลายน้ำต้นฉบับมาก

4.1 ผลการทดลองการส่งลายน้ำ

ผลการทดลองส่งลายน้ำบนสองโดเมน กราฟเปรียบเทียบการส่งลายน้ำโดเมนเดียวแบบต่าง ๆ โดยพิจารณาค่า PSNR ได้ดังตารางที่ 1

ตารางที่ 1: ค่า PSNR เปรียบเทียบการส่งลายน้ำสามประเภท

	สองโดเมน	โดเมนเวลา	โดเมนความถี่
5	47.9991	54.2841	54.9306
10	37.6222	44.5537	43.1749
20	30.1044	37.0359	32.9584
30	26.8046	33.7361	27.9321
50	23.7065	30.6380	22.7419

จะเห็นได้ว่าวิธีการที่นำเสนอได้อัตรา PSNR พอ ๆ กับเทคนิคการส่งลายน้ำบนโดเมนเดียว ซึ่งการที่ค่า PSNR ได้น้อยกว่าเทคนิคเหล่านั้นก็เพราะการส่งลายน้ำลงสองโดเมนส่งผล

ให้มีสัญญาณรบกวนมากขึ้น แต่ก็แลกมาด้วยความทนทานต่อการโจมตีรูปแบบต่าง ๆ นั้นเอง

เมื่อทำการฝังลายน้ำสำเร็จต่อไปจะทดลองการโจมตีภาพฝังลายน้ำ เพื่อวัดค่าความทนทานของลายน้ำหลังจากทำการถอดแล้วโดยวัดด้วยค่า NC

4.2 ผลการทดสอบความคงทนของลายน้ำ

การทดสอบความคงทนของลายน้ำกระทำโดยนำภาพที่ฝังลายน้ำแล้ว ทำการโจมตีขึ้นพื้นฐานและวัดค่า NC ซึ่งเป็นค่าเปรียบเทียบภาพต้นฉบับกับภาพที่มีสัญญาณรบกวน ยิ่งค่าเข้าใกล้ 100 ยิ่งเหมือนต้นฉบับ

การทดสอบแรกใช้การโจมตีด้วยการบีบอัดข้อมูลภาพ ซึ่งเป็นการโจมตีเชิงสัญญาณ ซึ่งเปรียบเทียบอัตราการบีบอัดที่ 90%, 70%, 50%, 30% และ 10% ตามลำดับ ตารางเปรียบเทียบความทนทานของลายน้ำจากการโจมตีด้วยการบีบอัดข้อมูลภาพ แสดงดังตารางที่ 2

ตารางที่ 2: การบีบอัดข้อมูลภาพขนาดต่าง ๆ

	90%	70%	50%	30%	10%
5	90.71	89.82	87.36	84.89	83.97
10	82.20	84.19	83.03	81.25	80.57
20	73.68	76.82	77.21	75.33	74.59
30	69.78	69.50	69.87	69.25	68.18
50	57.49	63.27	62.71	61.58	61.02

การทดสอบที่สองใช้การโจมตีด้วยการกรองความถี่ ซึ่งเป็นการโจมตีเชิงสัญญาณ ซึ่งเปรียบเทียบการกรองความถี่ 3 แบบด้วยกันคือ Average Filter, Gaussian Filter และ Sobel Filter ตามลำดับ ตารางเปรียบเทียบความทนทานของลายน้ำจากการโจมตีด้วยการกรองความถี่แสดงดังตารางที่ 3

ตารางที่ 3: การกรองความถี่รูปแบบต่าง ๆ

	Average	Gaussian	Sobel
5	85.6806	100	49.6751
10	86.1402	100	48.8888
20	86.2207	100	50.1117
30	86.2083	100	50.2131

50	86.1338	100	50.2711
----	---------	-----	---------

การทดสอบที่สามใช้การโจมตีด้วยการลดขนาดภาพ ซึ่งเป็นการโจมตีเชิงเรขาคณิต โดยนำภาพฝังลายน้ำมาลดขนาดภาพลงจาก 512x512 พิกเซล เป็น 32x32, 64x64, 128x128 และ 256x256 พิกเซล ตามลำดับ ตารางเปรียบเทียบความทนทานของลายน้ำจากการโจมตีด้วยการลดขนาดภาพ แสดงดังตารางที่ 4

ตารางที่ 4: การย่อภาพขนาดต่าง ๆ

	32x32	64x64	128x128	256x256
5	50.2256	50.4428	53.6047	63.6531
10	50.3207	50.865	54.0516	67.353
20	50.3828	50.6994	54.3807	68.3007
30	50.4056	50.8236	54.8504	65.9604
50	50.3435	50.9395	55.9719	64.0359

การทดสอบที่สี่ใช้การโจมตีด้วยการตัดขอบของภาพออก ซึ่งเป็นการโจมตีเชิงเรขาคณิต โดยนำภาพฝังลายน้ำตัดขอบออกเท่ากับ 5%, 10% และ 20% ของภาพ ตามลำดับ ตารางเปรียบเทียบความทนทานของลายน้ำจากการโจมตีด้วยการตัดขอบภาพแสดงดังตารางที่ 5

ตารางที่ 5: การตัดขอบภาพขนาดต่าง ๆ

	5%	10%	20%
5	58.6351	57.8529	56.0113
10	59.2083	58.4033	56.4955
20	59.349	58.5565	56.6113
30	59.3656	58.5751	56.63
50	59.3759	58.5834	56.6341

จากการทดสอบต่าง ๆ ทำให้พบว่าข้อดีของการฝังลายน้ำแต่ละแบบนี้แตกต่างกันไปตามแต่ละโดเมน ซึ่งการรวมข้อดีของทั้งสองอย่างด้วยกัน ก็มีข้อเสียเล็กน้อยนั่นคือสัญญาณรบกวนมากกว่า ค่า NC จึงอยู่ในระดับปานกลาง แต่เมื่อรูปถูกอัปโหลดบนอินเทอร์เน็ต การโจมตีหลายแบบพร้อมกันถือเป็นเรื่องที่

เกิดขึ้นทั่วไป ดังนั้นประสิทธิภาพของลายน้ำที่ดีจะต้องวัดด้วยการโจมตีหลายโดเมนพร้อม ๆ กัน

ดังนั้นการทดสอบที่ห้า จึงเป็นการ โจมตีทั้งสี่แบบพร้อมกัน นั่นคือ การบีบอัดข้อมูลภาพ การกรองความถี่ การลดขนาดภาพ และการตัดขอบภาพออก ผลลัพธ์แสดงดังตารางที่ 6

ตารางที่ 6: การ โจมตีสี่รูปแบบพร้อมกัน

	สองโดเมน	โดเมนเวลา	โดเมนความถี่
5	55.8228	50.0182	52.2423
10	61.4676	51.2065	53.8258
20	73.3368	52.2088	58.3077
30	81.4746	53.7001	61.8037
50	89.7634	56.7582	69.7054

จากการทดลองจะเห็นได้อย่างชัดเจนว่า การฝังลายน้ำบนทั้งสองโดเมนทำให้ได้ลายน้ำที่ทนทานต่อรูปแบบการโจมตีหลากหลายรูปแบบ มากกว่าการฝังลายน้ำบนโดเมนเดียว หลักฐานคือค่า NC ที่แสดงออกมานั่นเอง

5. สรุป

ในการทำภาพพิมพ์ลายน้ำดิจิทัลโดยทั่วไปมีหลักการสำคัญคือ หลังการใส่ลายน้ำ เทคนิคการฝังลายน้ำที่ดีนั้นไม่ควรจะทำให้คุณภาพลดลงมากจนเกินไป หากมีการ โจมตีภาพฝังลายน้ำโดยวิธีการประมวลผลสัญญาณแบบต่าง ๆ วิธีการเหล่านั้นต้องไม่ทำให้ลายน้ำสูญหายหรือถูกทำลายได้ นอกจากเจ้าของผลงานแล้ว ไม่ควรมีผู้ใดที่สามารถถอดสัญญาณลายน้ำออกจากรูปภาพที่ถูกฝังลายน้ำอยู่ได้

จากการศึกษาเทคนิคการฝังลายน้ำต่าง ๆ พบว่าแต่ละโดเมนนั้นมีข้อดีข้อเสียแตกต่างกันไป คือ โดเมนเวลาทนทานต่อการโจมตีเชิงเรขาคณิต เช่น การย่อภาพ ขยายภาพ ตัดขอบออก แต่สังเกตุดูเห็นลายน้ำได้ง่าย หรือ โดเมนความถี่ทนทานต่อการโจมตีเชิงสัญญาณ เช่น การบีบอัดข้อมูล การกรองความถี่ แต่จะสูญหายได้ง่ายเมื่อถูกโจมตีเชิงเรขาคณิต ซึ่งการรวมข้อดีของทั้งสองแบบเข้าด้วยกันทำให้ได้ลายน้ำที่ดีกว่า

สำหรับการฝังลายน้ำด้วยวิธีที่นำเสนอ นั้น มีข้อจำกัดคือสามารถใช้ได้กับภาพสี RGB เท่านั้น เนื่องจากต้องใช้ช่องสัญญาณสีจำนวนสองช่องด้วยกัน ทำให้ไม่สามารถใช้ภาพขาวดำได้ และใช้เวลาในการใส่ลายน้ำนานขึ้นเล็กน้อย เพราะต้องฝังลายน้ำในช่องสัญญาณสองช่องดังที่กล่าวมา

เอกสารอ้างอิง

- [1] M. George, J. Y. Chouinard and N. Georganas, "Digital Watermarking of Images and Video using Direct Sequence Spread Spectrum Techniques," *IEEE Trans. Computers in Cardiology*, 1999.
- [2] Thumrongrat Amornraksa, Pongsakorn Jirakulsawad and Bundit Thipakorn, "Digital Watermarking using Spread Spectrum Techniques," *วารสารวิจัยและพัฒนา มจร ฉบับที่ 2 เมษายน-มิถุนายน 2554*.
- [3] Yu-Pin Wang, Mei-Juan Chen and Po-Yuen Cheng, "Robust Image Watermark with Wavelet Transform and Spread," *IEEE Trans. Computers in Cardiology*, 2000.

การค้นพบกฎความสัมพันธ์จากระบบสนับสนุนการแก้ปัญหาการใช้งาน

Association Rule Discovery from Helpdesk Support System

ภัทราวุฒิ แสงศิริ (Patharawut Saengsiri)¹ ชนกชมน สุขศรี (Chanokthamon Suksri)² และพรนภา เจมส์ (Pornnapa James)³

กองเทคโนโลยีสารสนเทศและการสื่อสาร, ศูนย์ความรู้

สถาบันวิจัยวิทยาศาสตร์และเทคโนโลยีแห่งประเทศไทย

¹pvs@tistr.or.th, ²chanok_t@tistr.or.th, ³pornnapa@tistr.or.th

บทคัดย่อ

ปัจจุบันระบบสารสนเทศในองค์กรมีขนาดใหญ่มากขึ้น และมีความซับซ้อนทั้งในเรื่องของโครงสร้าง และขั้นตอนการปฏิบัติงาน ระบบสนับสนุนการแก้ปัญหาการใช้งาน (Helpdesk Support System) จึงเป็นระบบหนึ่งที่หลายองค์กรพัฒนาขึ้น เพื่อให้คำแนะนำและให้การสนับสนุนผู้ใช้งาน ให้สามารถปฏิบัติงานได้อย่างมีประสิทธิภาพมากขึ้น อย่างไรก็ตามการให้คำแนะนำและการสนับสนุนอาจไม่เพียงพอต่อการตอบสนองการแก้ปัญหา ดังนั้นงานวิจัยนี้จึงนำเสนอ การค้นหากฎความสัมพันธ์จากระบบสนับสนุนการแก้ปัญหาการใช้งาน เพื่อนำกฎที่ได้รับมาจัดวางแผนการฝึกอบรมให้กับผู้ใช้งาน และเพื่อป้องกันปัญหาที่จะเกิดขึ้นกับระบบในอนาคต ซึ่งผลลัพธ์แสดงการค้นพบกฎความสัมพันธ์ที่น่าสนใจ เช่น กลุ่มผู้ใช้งานรหัส S34 และ S53 เป็นกลุ่มผู้ร้องขอบริการด้านโปรแกรมมากที่สุดจากทุกหมวด ในขณะที่ผู้ให้บริการรหัส C1TD4 และ C1TD11 เป็นผู้ให้บริการมากที่สุดในทุกหมวด และให้บริการกับกลุ่มผู้ร้องขอบริการรหัส S26 ในเรื่องติดตั้งระบบ ERP มากที่สุด

คำสำคัญ: กฎความสัมพันธ์ การสนับสนุน ปัญหาการใช้งานระบบ

Abstract

Nowadays, information systems in organizations are growing up that lead to complexity of structure and operation procedure. Helpdesk Support System is developed to suggest and support user in many businesses to operate more efficiently. However, it is not enough for responding the problem. Therefore, this paper presents the discovery on

association rule from helpdesk support system that rule for planning of training schedule and preventing the problem in the future. The experiment shows the discovery on very interesting association rule such as user groups S34 and S35 that request service in all categories. As staff code C1TD4 and C1TD11 provide the most assistance in all service types, and install ERP system for user group S26 more than other user group.

Keywords: Association Rule, Supporting, System Problem, Helpdesk Support System

1. บทนำ

เมื่อระบบสารสนเทศถูกนำมาใช้งาน สิ่งที่น่าประหลาดใจหลังจากนั้นได้แก่ ความคาดหวังของผู้ใช้งาน ประสิทธิภาพของระบบที่ดี และการรักษาความปลอดภัย ซึ่งขั้นตอนการดูแลและการบำรุงรักษาระบบ (System maintenance) เป็นขั้นตอนถัดไปของกระบวนการพัฒนาระบบสารสนเทศ (System Development Life Cycle) ที่มีส่วนสำคัญในเรื่องดังกล่าว

ปัจจุบันวิธีการในการจัดการเรื่องของการดูแล และการบำรุงรักษาระบบได้ถูกนำมาพัฒนาเป็นระบบสารสนเทศ เรียกกันว่าระบบสนับสนุนการแก้ปัญหาการใช้งาน หรือ Helpdesk Support System ระบบดังกล่าวเริ่มได้รับการยอมรับอย่างกว้างขวาง ในหน่วยงานที่มีผู้ใช้งานเป็นจำนวนมาก และการสนับสนุนคือ การทำให้ผู้ใช้งานเรียนรู้คุณลักษณะของซอฟต์แวร์ และใช้ซอฟต์แวร์ให้เต็มประสิทธิภาพ รวมถึงช่วยให้ผู้ใช้งานค้นพบหนทางในการใช้

ซอฟต์แวร์ในการปฏิบัติงานให้มีประสิทธิภาพ [1] อย่างไรก็ตาม การให้คำแนะนำ และการสนับสนุนอาจไม่เพียงพอต่อการตอบสนองต่อการแก้ปัญหา ดังนั้นงานวิจัยนี้จึงนำเสนอ การค้นหา กฎความสัมพันธ์จากระบบสนับสนุนการแก้ปัญหาการใช้งาน เพื่อนำกฎที่ได้รับ มาจัดวางแผนการฝึกอบรมให้กับผู้ใช้งานและเพื่อป้องกันปัญหาที่จะเกิดขึ้นกับระบบในอนาคต

2. ทฤษฎีและวรรณกรรมที่เกี่ยวข้อง

2.1 ระบบสนับสนุนการแก้ปัญหาการใช้งาน

เป็นศูนย์กลางที่ช่วยรวบรวมข้อมูล และบริหารจัดการงาน บริการที่ได้รับการร้องขอจากผู้ให้บริการ ข้อมูลป้อนเข้า (Input) ที่ถูกนำเข้าระบบดังกล่าวจะถูกนำมา ประมวลผล เพื่อมอบหมายให้ผู้ให้บริการที่มีหน้าที่รับผิดชอบดำเนินการต่อไป โดยระบบมีหน้าที่ ความรับผิดชอบหลักดังนี้ [2]

1. ตอบปัญหาและข้อสงสัย (Answering questions) ได้แก่ การตอบปัญหาด้านเทคนิค และเฉพาะอุปกรณ์คอมพิวเตอร์ หรือ โปรแกรมที่อยู่ในความรับผิดชอบของหน่วยงาน
2. ให้บริการตามใบคำร้อง (Service a form request) ผู้ใช้งานสามารถร้องขอใช้บริการ โดยกรอกใบคำร้อง เช่น การขอให้ติดตั้งไดรเวอร์ (Driver install) หรือ การติดตั้งโปรแกรมป้องกันไวรัส (Antivirus program)
3. การปฏิบัติงานในกรณีเร่งด่วน (Handling an Emergency)
4. การแจ้งปัญหาเนื่องจากการขัดข้องของระบบให้กับผู้ใช้งานทราบ (Informing customer of system problems)
5. จัดทำรายงานให้กับผู้ดูแลระบบ (Reporting)

สำหรับช่องทางในการติดต่อสื่อสารกับระบบสนับสนุนการแก้ปัญหาการใช้งาน โดยทั่วไปมีหลายช่องทางได้แก่ โทรศัพท์ อีเมล อินทราเน็ต และแบบฟอร์ม เป็นต้น

โครงสร้างของระบบสนับสนุนการแก้ปัญหาการใช้งาน โดยทั่วไปประกอบด้วยโครงสร้างหลักที่สำคัญดังนี้ [3]

1. โครงสร้างแบบหลายชั้น คือ มีกลุ่มผู้ให้บริการระบบสนับสนุนแยกออกมากกว่าหนึ่งกลุ่ม เพื่อทำหน้าที่ให้บริการกับผู้ใช้งานแต่ละกลุ่ม อาทิ ชั้นแรกคือผู้ให้บริการที่ตอบและแก้ไข

ปัญหาทั่วไป ชั้นที่สอง คือ ผู้ให้บริการที่ตอบปัญหาด้านเทคนิคด้านหนึ่ง

2. โครงสร้างแบบรวมศูนย์หรือกระจายศูนย์ โครงสร้างลักษณะนี้เหมาะสำหรับองค์กรที่มีหน่วยงานกระจายอยู่ตามสถานที่ต่างกัน ดังนั้นผู้ให้บริการที่อยู่ตามสำนักงานจะมีความเหมาะสมในการปฏิบัติงานทั้งในเรื่องของความสะดวกในการเดินทางและความรวดเร็วในการให้บริการ

นอกจากนี้ยังมีงานวิจัยที่เกี่ยวข้องการพัฒนา ระบบสนับสนุนการแก้ปัญหาการใช้งาน เช่น [4] ได้พัฒนาระบบสืบค้นข้อมูลเพื่อช่วยสนับสนุนการแก้ไขปัญหาในระบบคอมพิวเตอร์สำหรับกรณีที่ต้องการช่วยจัดเก็บองค์ความรู้ที่ได้จากเจ้าหน้าที่ที่มีประสบการณ์ในการแก้ไขปัญหา ทำให้นำไปสู่การตัดสินใจที่ถูกต้องแม่นยำ และมีความผิดพลาดน้อยที่สุด อย่างไรก็ตามระบบดังกล่าวมีข้อเสียที่ไม่รองรับการมอบหมายงานในลักษณะโครงสร้างแบบหลายชั้น หรือ [5] ได้พัฒนาระบบสนับสนุนการแก้ปัญหาการใช้งาน ADSL สำหรับเจ้าหน้าที่ Helpdesk ภูมิศึกษา บริษัท ทีโอที จำกัด (มหาชน) ซึ่งเป็นระบบเก็บคำถาม และคำตอบเพื่อสนับสนุนการทำงาน สาเหตุเนื่องจากบริษัทมีลูกค้าจำนวนมาก ดังนั้นกลยุทธ์การรักษาลูกค้าโดยการให้บริการหลังการขายจึงมีความสำคัญสูงมาก การรอแต่เพียงผู้เชี่ยวชาญอาจไม่สามารถแก้ไขปัญหาได้ทันทั่วทั้ง และที่สำคัญต้องสามารถให้บริการได้ตลอด 24 ชั่วโมง อย่างไรก็ตามระบบนี้ยังมีข้อเสียที่สำคัญ เกี่ยวกับจำนวนคำถามที่ต้องการเพิ่มให้กับระบบให้มากยิ่งขึ้น อีกทั้งยังไม่สามารถรองรับกับบริการอื่นของบริษัทได้ สำหรับงานวิจัยที่น่ากฎความสัมพันธ์ไปใช้เป็นเครื่องมือในการวิจัย ได้แก่ [6] นำเสนอการใช้กฎความสัมพันธ์เนื่องจากปัญหานักพัฒนามักตั้งชื่อเมธอดของการพัฒนาโปรแกรมเชิงวัตถุไม่สอดคล้องกัน โดยงานวิจัยนี้สามารถแนะนำคำกริยาที่ใช้แทนชื่อเมธอดได้ อย่างไรก็ตามมีข้อเสียที่สำคัญคือ กฎที่ได้มายังมีข้อจำกัดแค่คำกริยาเท่านั้น [7] นำเสนอระบบแนะนำสินค้าด้วยวิธี Collaborative Filtering และวิธี item-based Collaborative Filtering ร่วมกับการแนะนำสินค้าที่เกี่ยวข้องโดยใช้ Association Rules อย่างไรก็ตาม งานวิจัยนี้ยังขาดความสนใจในเรื่องการค้นหาสินค้าที่เกี่ยวข้องกับแต่ละผู้ซื้อแต่ละคน โดยมุ่งไปที่สินค้าที่เกี่ยวข้องกับกลุ่มผู้ซื้อเป็นหลัก ซึ่งอาจจะไม่ได้ผลดีเสมอไป [8] นำเสนอ

งานวิจัยเกี่ยวกับการทำเหมืองข้อมูลในโรงพยาบาล โดยใช้เทคนิค กฎความสัมพันธ์ ทำให้ค้นพบกฎที่เกี่ยวข้องระหว่างโรค และปัจจัยที่สัมพันธ์กัน อย่างไรก็ตามงานวิจัยนี้ยังต้องการศึกษาเพิ่มเติม เพื่อให้สามารถแสดงกฎที่มีความสำคัญแต่มีค่านับสนับสนุนน้อย ทำให้ไม่สามารถแสดงกฎดังกล่าวได้

2.2 กฎความสัมพันธ์ (Association Rule)

กฎความสัมพันธ์ เป็นวิธีการหนึ่งของการทำเหมืองข้อมูล โดยมีเป้าหมายเพื่อใช้วิเคราะห์หาความสัมพันธ์ที่ซ่อนอยู่ในกลุ่มข้อมูลขนาดใหญ่ [9] ทั้งนี้เพื่อนำไปพยากรณ์เหตุการณ์ต่าง ๆ เช่น การซื้อสินค้าในซูเปอร์มาร์เก็ต หรือ Market Basket Analysis ที่เป็นการค้นหาความสัมพันธ์ของรายการสินค้าที่ถูกซื้อ และมีประโยชน์ในการจัดเรียงสินค้าบนชั้นของร้านค้า เพื่ออำนวยความสะดวกให้กับลูกค้าได้ดียิ่งขึ้น รวมถึงสามารถนำไปใช้ป็นเครื่องมือในการส่งเสริมการขายสินค้า

กฎความสัมพันธ์เขียนอยู่ในรูป $x \rightarrow y$ โดยที่ $x, y \subseteq I$

และ $x \cap y = \emptyset$, $x \rightarrow y$ [Support = S, Confidence] หมายถึงกฎ $x \rightarrow y$ ที่มีค่านับสนับสนุนเท่ากับ S ในเซตของรายการ (Transaction) จากฐานข้อมูล และ D หมายถึง S% ของจำนวนรายการทั้งหมดประกอบด้วย *itemsets* x และ y (เขียนแทนด้วย $x \cup y$) และมีค่าความเชื่อมั่นเท่ากับ C ถ้า C% ของรายการที่ประกอบด้วย *itemsets* x แล้วรายการเหล่านั้นมี *itemsets* y ด้วย [10]

- Itemset คือ ความสัมพันธ์ของข้อมูลที่หาได้ โดย Itemset ประกอบด้วย item ที่ k-itemsets เช่น 3-itemsets {Milk, Bread, Diaper}
- Support count คือ ความถี่ของการปรากฏ itemset เช่น {Milk, Bread, Diaper}
- Support คือ ค่านับสนับสนุน หรือ ค่าความจริงของจำนวนรายการที่มี itemset อยู่ ทั้งนี้เป็นการบ่งบอกถึงคุณภาพของกฎ อาทิ s({Milk, Bread, Diaper}) ดังแสดงในสมการ 1

$$Support(x) = \frac{\text{จำนวน Transaction ที่ประกอบด้วย Itemset } x}{\text{จำนวน Transaction ทั้งหมดในฐานข้อมูล}} \quad (1)$$

- Frequent Itemset คือ itemset ที่มีค่านับสนับสนุนสูงกว่าหรือเท่ากับค่านับสนับสนุนขั้นต่ำ (Minimum support) เช่น กำหนดค่า minimum support = 0.5 ดังนั้น itemset ที่สามารถเป็น frequent itemset ได้ต้องมีค่านับสนับสนุนสูงกว่าหรือเท่ากับ 0.5
- ค่าความเชื่อมั่น (Confidence value) เป็นค่าแสดงความเป็นจริงของรายการเปรียบเทียบกับจำนวนรายการที่มีทั้งหมด สามารถคำนวณได้ดังสมการ 2

$$Confidence(x \rightarrow y) = \frac{Support(xy)}{Support(x)} \quad (2)$$

2.3 อัลกอริทึม Apriori

ขั้นตอนวิธีการค้นหาความสัมพันธ์ที่รู้จักกันดีได้แก่ Apriori อัลกอริทึม ซึ่งใช้หลักการของการคำนวณหาความถี่ของ itemsets ที่เกิดขึ้นพร้อมกันในฐานะข้อมูล ซึ่งมีรูปแบบดังนี้ {Item1, Item2} -> {item3}; Support, Confidence

เงื่อนไขสำคัญของการใช้งานอัลกอริทึมนี้ คือ การกำหนดค่า Support และค่า Confidence ขั้นต่ำโดยผู้ใช้งานสามารถกำหนดขึ้นเอง หรืออาจให้ผู้เชี่ยวชาญเป็นผู้กำหนดขึ้น [11] ทั้งนี้กฎความสัมพันธ์ที่ค้นพบได้นั้นต้องมีทั้งสองดังกล่าวสูงกว่าค่าที่กำหนดไว้

3. วิธีการดำเนินการวิจัย

3.1 การเตรียมข้อมูล

งานวิจัยนี้ใช้ข้อมูลจากฐานข้อมูล Helpdesk Support System ของสถาบันวิจัยวิทยาศาสตร์และเทคโนโลยีแห่งประเทศไทย (วว.) ซึ่งมีข้อมูลจำนวน 2 ชุด คือ 1) มีจำนวน 62,499 ระเบียน มีมิติข้อมูลทั้งสิ้น 6 มิติข้อมูล 2) มีจำนวน 12,080 ระเบียน มีมิติข้อมูลจำนวนทั้งสิ้น 4 มิติข้อมูล จากนั้นทำการแปลงข้อมูลให้อยู่ในรูปแบบข้อความ (Text) แทนที่ชื่อตำแหน่งด้วยรหัส "S" แทนที่ชื่อผู้ให้บริการดูแลบำรุงรักษาคอมพิวเตอร์ด้วยรหัส "CITD" ซึ่งมีค่าตั้งแต่ CITD1 ถึง CITD11

3.2 การสร้างโมเดลโดยใช้เทคนิคกฎความสัมพันธ์

สำหรับการสร้าง โมเดลโดยใช้เทคนิคกฎความสัมพันธ์ ซึ่งงานวิจัยนี้ เลือกเทคนิคแบบ Apriori และใช้โปรแกรม Weka เวอร์ชัน 3.6 โดยทำการทดลองสร้างโมเดลกับโปรแกรม Weka

ตารางที่ 1: มิติข้อมูลที่ใช้งานวิจัย

ลำดับที่	ข้อมูลชุดที่ 1	ข้อมูลชุดที่ 2
1.	Position_ID	Customer_ID
2.	Location	Time1
3.	Date	Name_Eng
4.	Time	Response_name
5.	Service_Eng	-
6.	Response_name	-

ตารางที่ 2: ตัวอย่างชุดข้อมูลที่ 1

Position_ID	Location	Date	Time	Service_Eng	Response_name
S34	L1	2-Jan-13	T9	Install program	CITD4
S28	L1	2-Jan-13	T16	Install computer meeting room	CITD11
S53	L1	3-Jan-13	T9	Borrow peripherals	CITD11
S42	L1	2-Jan-13	T10	TISTR website problem	CITD1
S27	L1	3-Jan-13	T10	OS Configuration	CITD8

ตารางที่ 3: ตัวอย่างชุดข้อมูลที่ 2

Customer_ID	Time	Name_Eng	Response_name
C774	T15	Install_program	citd6
C185	T15	Install_program	citd3
C738	T09	Share_printer	citd3
C254	T10	Install_driver	citd6
C351	T17	troubleshooting	citd6
C105	T10	Virus	citd11
C36	T13	Wiring_LAN	citd6
C196	T09	Map_drive	citd2

จากนั้นได้ทำการทดลองเปลี่ยนค่าสนับสนุน และค่าความเชื่อมั่นในระดับที่แตกต่างกัน ทั้งนี้เนื่องจากกลุ่มข้อมูลที่นำมาทดลองหา

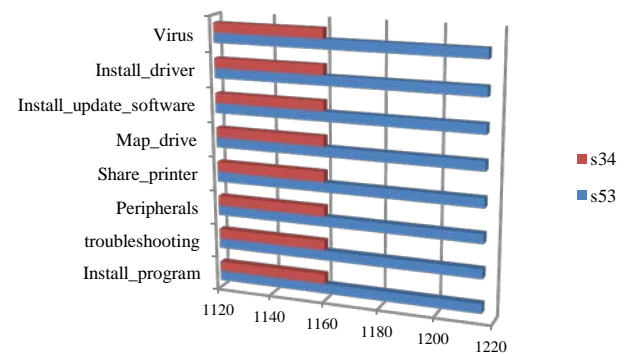
ความสัมพันธ์มีรายการ (ระเบียบข้อมูล) แตกต่างจากกลุ่มข้อมูล

โดยทั่วไป ดังตัวอย่างชุดข้อมูลที่ 1 และ 2 แสดงในตารางที่ 2 และ 3

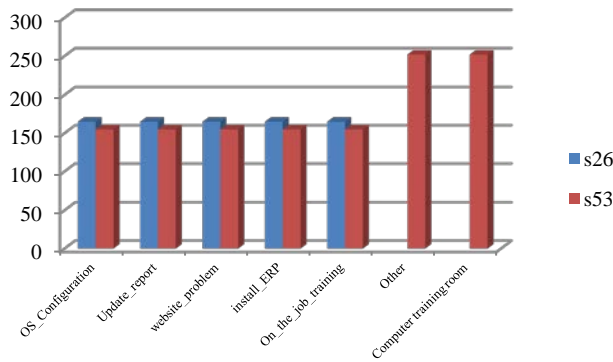
ดังนั้นงานวิจัยนี้ จึงได้ทำการจากการทดลองการกำหนดค่าสนับสนุนและค่าความเชื่อมั่นที่แตกต่างกัน เพื่อคัดเลือกค่าสนับสนุนและค่าความเชื่อมั่นที่เหมาะสม สำหรับการคัดแยกกลุ่มข้อมูลชุดที่ 1 และชุดที่ 2 พบว่า การกำหนดค่าความเชื่อมั่น และค่าสนับสนุนมีค่าที่เหมาะสมที่สามารถค้นหาความสัมพันธ์ได้ คือ ค่าสนับสนุนที่ 0.009 และค่าความเชื่อมั่นที่ 0.09

3.3 การค้นหาความสัมพันธ์จากกลุ่มข้อมูลชุดที่ 1

วัตถุประสงค์ของขั้นตอนนี้คือ ต้องการทราบกฎความสัมพันธ์ที่ถูกละทิ้งอยู่ในกลุ่มข้อมูลชุดที่ 1 โดยกำหนดค่าสนับสนุนและค่าความเชื่อมั่นตามข้อ 3.2 ผลลัพธ์ของกฎความสัมพันธ์ที่ค้นพบได้โดยใช้อัลกอริทึม Apriori จากกลุ่มข้อมูลชุดที่ 1 ได้แก่ 1) กลุ่มผู้ใช้งานที่มีรหัส S34 และ S53 เป็นผู้ร้องขอบริการด้าน Virus, Install driver, Install update software, Map driver, Share printer, Peripherals, Troubleshooting และ Install program มากที่สุด 2) กลุ่มผู้ใช้งานที่มีรหัส S26 และ S53 เป็นผู้ร้องขอบริการด้าน OS Configuration, Update report, Website problem, Install ERP, On the job training, Computer training room และอื่น ๆ มากที่สุด

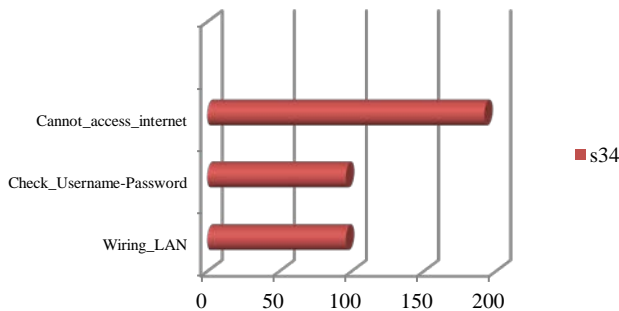


ภาพที่ 1: บริการที่ผู้ใช้งานกลุ่ม S34 และ S53 ร้องขอบริการ

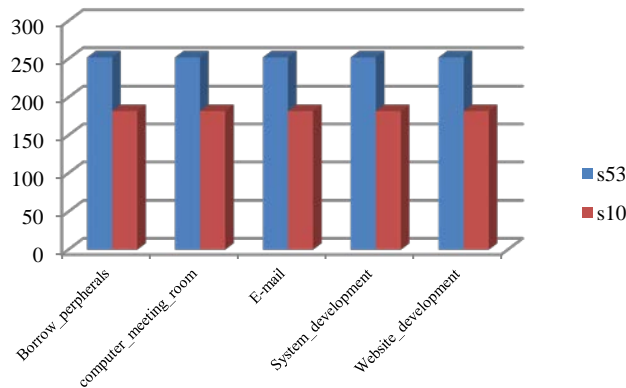


ภาพที่ 2: บริการที่ผู้ใช้งานกลุ่ม S26 และ S53 ร้องขอบริการ

3) กลุ่มผู้ใช้งานที่มีรหัส S34 เป็นผู้ร้องขอบริการการแก้ไขปัญหาเรื่อง access internet, Check username password และ Wiring LAN มากที่สุด ดังแสดงในภาพที่ 3 และกลุ่มผู้ใช้งานที่มีรหัส S53 และ S10 เป็นผู้ร้องขอบริการด้าน Borrow peripherals, Computer meeting room, E-mail, System development และ Website development มากที่สุด ดังแสดงในภาพที่ 4



ภาพที่ 3: บริการที่ผู้ใช้งานกลุ่ม S34 ร้องขอบริการแก้ไขปัญหา



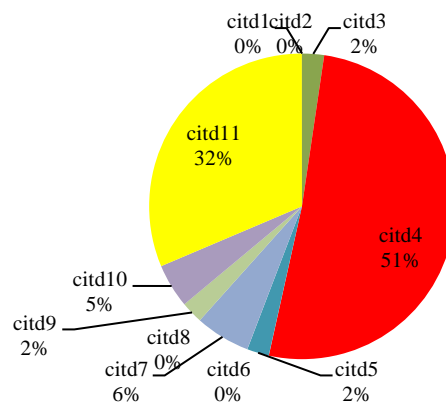
ภาพที่ 4: บริการที่ผู้ใช้งานกลุ่ม S53 และ S10 ร้องขอบริการ

3.4 การค้นหาความสัมพันธ์จากกลุ่มข้อมูลชุดที่ 2

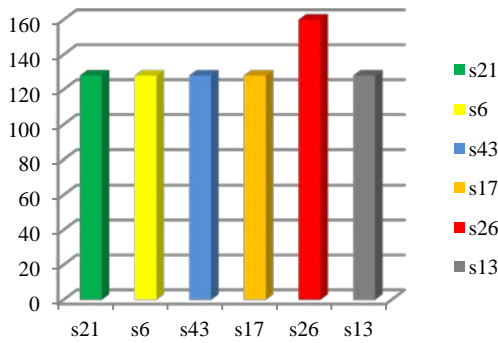
สำหรับกลุ่มข้อมูลชุดที่ 2 เพื่อนำมาค้นหาความสัมพันธ์ทำให้ค้นพบความสัมพันธ์ ที่มีผลลัพธ์เป็นสาระสำคัญของการให้บริการ ได้แก่ 1) รหัสผู้ให้บริการที่ออกปฏิบัติงานให้กับหน่วยงานซึ่งแสดงเป็นเปอร์เซ็นต์ (Percentages) โดยมีรหัสของผู้ให้บริการสูงสุดคือ CITD4 51% และ CITD11 32% ตามลำดับดังแสดงในภาพที่ 5 และ 2) กลุ่มของผู้ใช้บริการที่ได้รับการบริการสูงสุดจาก Helpdesk Support System คือ กลุ่มรหัส S26 ดังแสดงในภาพที่ 6

3.5 ตัวอย่างกฎความสัมพันธ์ที่ค้นพบ

เมื่อทำการค้นหาความสัมพันธ์จากกลุ่มข้อมูลชุดที่ 1 และ 2 โดยการกำหนดค่าสนับสนุนและค่าความเชื่อมั่นตามที่ได้ทดลองในข้อ 3.2 และนำกฎที่ได้รับมาคัดเลือกโดยตัดกฎที่มีความซ้ำซ้อนกันออก ผลลัพธ์ที่ได้จากการทดลอง เช่น Service_Eng=OS Configuration 480 ==> Position_ID=S26 165 conf:(0.34), Service_Eng=Update report 480 ==> Position_ID=S26 165 conf:(0.34) และ Service_Eng= Install ERP 480 ==> Position_ID=S53 155 conf:(0.32) เป็นต้น



ภาพที่ 5: รหัสผู้ให้บริการและเปอร์เซ็นต์การให้บริการ



ภาพที่ 6: กลุ่มผู้ใช้งานที่ได้รับบริการสูงสุดจาก Helpdesk Support System

4. สรุปและอภิปรายผล

จากการทดลองเพื่อค้นหากฎความสัมพันธ์ โดยใช้ข้อมูล 2 ชุดจากระบบ Helpdesk Support System เป็นข้อมูลป้อนเข้าสำหรับอัลกอริทึม Apriori พบว่ากฎความสัมพันธ์สามารถเปิดเผยถึงข้อมูลที่ซ่อนอยู่ในฐานข้อมูลขนาดใหญ่ของกลุ่มข้อมูลชุดที่ 1 ดังนี้ กลุ่มผู้ใช้งานที่มีรหัส S34 และ S53 เป็นผู้ร้องขอบริการด้านโปรแกรม Virus, Install driver, Install update software, Map driver, Share printer, Peripherals, Troubleshooting และ Install program มากที่สุดมีจำนวนครั้ง 1,216 และ 1,160 ตามลำดับ จากกลุ่มผู้ใช้งานทั้งหมด ในขณะที่การขอบริการด้าน OS Configuration, Update report, Website problem, Install ERP และ On the job training มีจำนวนอย่างละ 155 ครั้งและ Computer training room และอื่น ๆ 252 ครั้ง ตามลำดับ ซึ่งเป็นการร้องขอจากกลุ่มผู้ใช้งานรหัส S26 และ S53 มากที่สุด ในส่วนของการร้องขอบริการเพื่อแก้ไขปัญหา access internet มีจำนวน 192 ครั้ง Check username password และ Wiring LAN มีจำนวนอย่างละ 96 ครั้ง โดยถูกร้องขอจากกลุ่มผู้ใช้งาน S34 มากที่สุด นอกจากนี้ผู้ร้องขอบริการด้าน Borrow peripherals, Computer meeting room, E-mail, System development และ Website development โดยกลุ่มผู้ขอใช้บริการมากที่สุดคือรหัส S53 มีจำนวน 252 ครั้งและ S10 จำนวน 182 ครั้ง

สำหรับกฎความสัมพันธ์ที่ได้จากกลุ่มข้อมูลชุดที่ 2 ทำให้สามารถเปิดเผยความสัมพันธ์ในเรื่องของรหัสผู้ให้บริการสูงสุดคือ CITD4 51% และ CITD11 32% ตามลำดับจากผู้ให้บริการทั้งหมด โดยกลุ่มของผู้ใช้บริการที่ขอรับบริการสูงสุดจาก Helpdesk

Support System คือ กลุ่มรหัส S26 มีจำนวน 160 ครั้ง ในขณะที่กลุ่มอื่นที่เหลือมีจำนวนการขอรับบริการเท่ากัน คือ มีจำนวน 128 ครั้ง

สำหรับงานวิจัยที่จะทำต่อไปในอนาคต คือการสร้างอัลกอริทึมเพื่อการค้นหากฎความสัมพันธ์ สำหรับข้อมูลที่มีค่าสนับสนุนและค่าความเชื่อมั่นต่ำ เพื่อให้สามารถใช้งานได้กับกลุ่มข้อมูลที่มีความหลากหลายมากยิ่งขึ้น

เอกสารอ้างอิง

- [1] Gary B. Shelly and Harry J. Rosenblatt, "System Analysis and Design Ninth Edition," *Course Technology, Cengage Learning*, 2012.
- [2] อริ อยู่ภู, "ระบบช่วยเหลือและแก้ปัญหาของการใช้งานโปรแกรมบริหารงานบุคคล," สารนิพนธ์ วิทยาศาสตร์มหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2554.
- [3] สุภัทรา พรหมสาขา ณ สกลนคร, "ระบบสนับสนุนการวิเคราะห์ปัญหาสำหรับเจ้าหน้าที่ Helpdesk โดยเทคโนโลยีออนไลน์," สารนิพนธ์ วิทยาศาสตร์มหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2555.
- [4] วลัยพร พุทธิปรังค์, "การพัฒนาเว็บสืบค้นข้อมูลเพื่อช่วยสนับสนุนการแก้ปัญหาระบบงานคอมพิวเตอร์สำหรับกรมที่ดิน," สารนิพนธ์ วิทยาศาสตร์มหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2551.
- [5] ละออง สุนทรแสง, "ระบบสนับสนุนการแก้ปัญหาการใช้งาน ADSL สำหรับเจ้าหน้าที่ Helpdesk กรณีศึกษา บริษัท ทีโอที จำกัด มหาชน," สารนิพนธ์ วิทยาศาสตร์มหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2552.
- [6] Kashiwabara, Y., Y. Onizuka, "Recommending verbs for rename method using association rule mining," *IEEE Conference on Software Maintenance, Reengineering and Reverse Engineering (CSMR-WCRE)*, pp. 323-327, 2014.
- [7] Choi, Y. K. and S. K. Kim, "An auxiliary recommendation system for repetitively purchasing items in E-commerce," *International Conference on Big Data and Smart Computing*

(*BIGCOMP*), pp. 96-98, 2014.

- [8] Shunmin, W., "The Study of Association Rule Mining Technology in Hospital Information System Analysis," *5th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, pp. 245-248, 2013.
- [9] Pang-Ning Tan, Michael Steinbach, and Vipin Kumar. "Introduction to Data Mining." *Addison Wesley*, pp. 158-164, 2006.
- [10] สมโภช พรหมแก้ว. "เว็บไซต์แนะนำเพื่อนเก่า โดยใช้เทคนิค Apriori." *สารนิพนธ์ วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*. 2550.
- [11] กฤษณากร กิ่งอุบล. "การใช้เกณฑ์ความต่างลำดับในการปรับปรุงคุณภาพสัมพันธ์จำแนกประเภทข้อมูล." *วิทยานิพนธ์ วิศวกรรมศาสตรมหาบัณฑิต (วิศวกรรมคอมพิวเตอร์) มหาวิทยาลัยเกษตรศาสตร์*. 2552.

การประยุกต์ใช้ขั้นตอนวิธีเชิงพันธุกรรมในการกำหนดกำลังทางเรือ: กรณีศึกษาภารกิจการช่วยเหลือด้านมนุษยธรรมและการบรรเทาภัยพิบัติ

Application of Genetic Algorithm in Maritime Force Composition: Case Study in Humanitarian Assistance and Disaster Relief Operations

ธนภม บุญแสนพล(Thanapon Boonsanpol)¹ และณรงค์เดช กิระติพรานนท์ (Narongdech Keeratipranon)²

¹ภาควิชาวิศวกรรมคอมพิวเตอร์และโทรคมนาคม คณะวิศวกรรมศาสตร์ มหาวิทยาลัยธุรกิจบัณฑิตย์

²ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยธุรกิจบัณฑิตย์

thanapon.bsp@hotmail.com, narongdech.ken@dpu.ac.th

บทคัดย่อ

การกำหนดกำลังทางเรือที่เหมาะสมด้วยตัวบุคคลเป็นปัญหาที่มีความซับซ้อนไม่สามารถกระทำได้ในระยะเวลาอันสั้น เนื่องจากเรือในกองทัพมีหลายประเภทและมีขีดความสามารถหรือสมรรถนะของเรือที่แตกต่างกัน อีกทั้งเรือบางประเภทยังสามารถตอบสนองภารกิจได้หลากหลายรูปแบบ งานวิจัยนี้จึงนำขั้นตอนวิธีเชิงพันธุกรรมมาประยุกต์ใช้เพื่อกำหนดกองกำลังทางเรือที่เหมาะสม สามารถตอบสนองภารกิจตามที่ต้องการได้ ทั้งนี้เพื่อเป็นการลดขั้นตอนที่ซับซ้อนในการค้นหากองกำลังทางเรือที่เหมาะสมด้วยตัวบุคคล อีกทั้งขั้นตอนวิธีเชิงพันธุกรรมยังสามารถให้คำตอบได้หลากหลายคำตอบที่สามารถที่จะนำมาประกอบการพิจารณาการตัดสินใจในการเลือกกองกำลังทางเรือได้อีกด้วย

คำสำคัญ: ขั้นตอนวิธีเชิงพันธุกรรม กำหนดกำลังทางเรือ

ภารกิจด้านมนุษยธรรม

Abstract

Maritime force composition is a complicated process especially when manually organizes the fleet. There are many factors involve in the selection process. The Navy has many type of warships each one has different characteristic and performance. Furthermore, there are many multi-function warships in the Royal Thai Navy. This research has investigated and successfully applied a genetic algorithm to the maritime force composition for a humanitarian assistance and disaster relief operations (HADR). The experiment shows a positive fleet result

comparable to manually force composition done by navy. With GA, there are many proposed fleet solutions for further select by a high rank commander.

Keyword: Genetic algorithm, Maritime force composition, Humanitarian Assistance and Disaster Relief Operations (HADR).

1. บทนำ

ทศวรรษที่ผ่านมาเป็นช่วงเวลาที่ยกยิบยิบขนาดใหญ่เกิดขึ้นเป็นจำนวนมาก สร้างความเสียหายทั้งชีวิต ทรัพย์สิน ระบบเศรษฐกิจ สังคม และสิ่งแวดล้อมเป็นมูลค่ามหาศาล ประชากรที่ได้รับผลกระทบจากภัยพิบัติทั่วโลกมีจำนวนถึง 4,000 ล้านคน จำนวนผู้เสียชีวิตประมาณ 2 ล้านคน และในอนาคตการเกิดภัยพิบัติมีแนวโน้มที่จะทวีความรุนแรงมากขึ้น เนื่องจากการเปลี่ยนแปลงทางสภาพภูมิอากาศ และภาวะโลกร้อน [1] ดังนั้นเมื่อมีภัยพิบัติเกิดขึ้น ผู้ประสบภัยอาจมีความต้องการในการร้องขอความช่วยเหลือต่างๆ จากกองทัพเรือ เช่น การอพยพเคลื่อนย้ายผู้ประสบภัย การกู้ภัย และการลำเลียงอาหาร เป็นต้น ซึ่งจากความต้องการดังกล่าวนี้ ทางกองทัพเรือจำเป็นต้องกำหนดกองกำลังทางเรือที่มีความเหมาะสม เพื่อตอบสนองต่อความต้องการในการเข้าช่วยเหลือจากสถานการณ์ที่เกิดขึ้น

ในปัจจุบันการกำหนดกำลังทางเรือจะใช้วิธีการกำหนดกำลังทางเรือตามภารกิจและขีดความสามารถของเรือ ซึ่งจะถูกจัดสรรโดยเสนาธิการทหารเรือ หรือด้วยตัวบุคคล และวิธีการในการจัดสรรกองกำลังเรือนี้จะมีวิธีการ 3 ขั้นตอน คือ 1.) ประเมินความต้องการจากกองกำลังทางเรือที่เหมาะสมสำหรับสถานการณ์ 2.) ประเมินขีดความสามารถของเรือในแต่ละประเภทที่พร้อมใช้งาน และจัดสรรเรือที่มีขีดความสามารถ

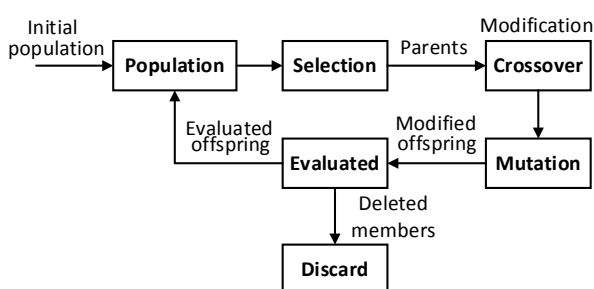
สอดคล้องต่อความต้องการ และ 3.) กำหนดเงื่อนไขหรือข้อจำกัดเพิ่มเติม เช่น ค่าใช้จ่ายของเรือในแต่ละประเภท ซึ่งจากขั้นตอนดังกล่าวนี้เป็นขั้นตอนที่จะต้องประมวลผลข้อมูลจำนวนมากในการค้นหาของกำลังทางเรือที่เหมาะสม ด้วยการวิเคราะห์ขีดความสามารถของเรือที่สามารถตอบสนองต่อภารกิจได้ ซึ่งไม่สามารถกระทำได้ในระยะเวลาอันสั้นด้วยตัวบุคคล เนื่องจากเรือในกองทัพเรือมีจำนวนประมาณกว่า 324 ลำ และเรือแต่ละลำถูกสังกัดไปตามหน่วยงานทหารเรือหลากหลายที่ในประเทศไทย อีกทั้งสมรรถนะของเรือในแต่ละประเภทมีขีดความสามารถที่หลากหลาย และสามารถตอบสนองต่อภารกิจได้หลากหลายรูปแบบ

ดังนั้นในงานวิจัยนี้จึงมีวัตถุประสงค์เพื่อพัฒนาระบบที่สามารถจัดสรรกองกำลังทางเรือที่เหมาะสมได้ ด้วยวิธีเชิงพันธุกรรมมาวิเคราะห์ข้อมูลและค้นหาคำตอบ เพื่อลดขั้นตอนที่ซับซ้อนในการค้นหาของกำลังทางเรือที่เหมาะสมด้วยตัวบุคคล

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 แนวทฤษฎีขั้นตอนวิธีเชิงพันธุกรรม

ขั้นตอนวิธีเชิงพันธุกรรม (Genetic Algorithm: GA) เป็นเทคนิคทางปัญญาประดิษฐ์อย่างหนึ่งที่ใช้ในการค้นหาคำตอบ โดยอาศัยหลักการของศาสตร์ทางด้านชีววิทยา ในเรื่องของการคัดเลือกแบบธรรมชาติ และหลักการทางสายพันธุ์ ซึ่งจัดได้ว่าเป็นการคำนวณแบบเป็นเชิงวิวัฒนาการในการค้นหาคำตอบ โดยสิ่งมีชีวิตที่เหมาะสมกับสภาพแวดล้อมในขณะนั้นจะสามารถมีชีวิตรอดต่อไปด้วยการกระทำโดยทางพันธุกรรม เช่น การสืบพันธุ์, การกลายพันธุ์ และการแลกเปลี่ยนยีน ซึ่งมีขั้นตอนการทำงานแสดงดังภาพที่ 1



ภาพที่ 1: กระบวนการทำงานขั้นตอนวิธีเชิงพันธุกรรม

กระบวนการค้นหาคำตอบของขั้นตอนวิธีเชิงพันธุกรรม จะเริ่มต้นโดยการกำหนดกลุ่มของคำตอบให้เป็นประชากรเริ่มต้น แล้วเข้าสู่กระบวนการคัดเลือกพันธุกรรมตามธรรมชาติ เพื่อปรับปรุงประชากรในรุ่นถัดไปให้มีความเหมาะสมมากยิ่งขึ้น หรืออยู่รอดเพื่อใช้ในการสร้างประชากรในรุ่นถัดไป โดยจะเริ่มจากการคัดเลือก (Selection) คำตอบที่เหมาะสมในรุ่นปัจจุบัน เพื่อเป็นพ่อแม่พันธุ์แล้วจึงทำการผสมคำตอบ (Crossover) ระหว่างพ่อพันธุ์และแม่พันธุ์ จากนั้นจะทำการคัดลอกส่วนของคำตอบเพื่อไปเข้าสู่กระบวนการกลายพันธุ์ (Mutation) แล้วจึงนำกลุ่มของคำตอบใหม่มาทำการประเมิน (Evaluate) เพื่อที่จะทำการคัดเลือกประชากรตัวที่เหมาะสมมาเป็นประชากรในรุ่นถัดไป ซึ่งกระบวนการนี้จะทำวนซ้ำไปเรื่อยๆจนกว่าจะได้คำตอบที่น่าพอใจ หรือจนกว่าประชากรรุ่นถัดไปไม่มีการพัฒนาให้ดีขึ้น

2.2 งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องกับการใช้ขั้นตอนวิธีเชิงพันธุกรรมในทางด้านต่างๆ มีตัวอย่างดังต่อไปนี้

วนิดา [2] ได้นำเทคนิคขั้นตอนวิธีเชิงพันธุกรรม มาพัฒนาระบบการจัดตารางสอนสำหรับนักศึกษา เพื่อให้ได้ตารางการเรียนสอนที่เหมาะสม โดยขั้นตอนวิธีเชิงพันธุกรรมจะช่วยลดระยะเวลาและข้อผิดพลาดที่เกิดขึ้นในการจัดตารางเวลา อาทิ เช่น การซ้ำซ้อนของวัน-เวลาเรียนของนักศึกษาที่ลงทะเบียน, การซ้ำซ้อนของวัน-เวลาสอนของอาจารย์ผู้สอน และการซ้ำซ้อนของวัน-เวลาในการใช้ห้องเรียน

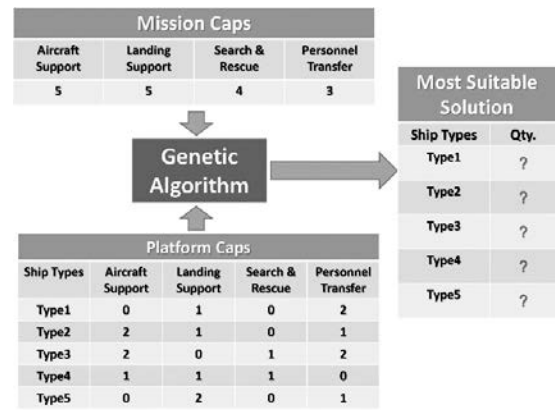
อุดม [3] ได้พัฒนาระบบโดยใช้ขั้นตอนวิธีเชิงพันธุกรรม สำหรับปัญหาการจัดเส้นทางการเดินทาง เพื่อค้นหาเส้นทางที่เหมาะสมในการรับส่งพนักงานระหว่างบริษัทและจุดรับส่งพนักงานโดยมีค่าใช้จ่ายที่ต่ำที่สุดหรือมีระยะทางรวมน้อยที่สุด ทั้งนี้เพื่อลดค่าใช้จ่ายในการเดินทางให้น้อยที่สุด

Chomtip, Verachad, Satheanpong and Nannaphat [4] ได้พัฒนาระบบการรู้จำลายมือเขียนตัวอักษรภาษาไทย ด้วยขั้นตอนวิธีเชิงพันธุกรรม เพื่อค้นหาว่าลายมือที่เขียนขึ้นเป็นตัวอักษรใดในภาษาไทย

กรรณิกา [5] ได้นำเสนอการเอาขั้นตอนวิธีเชิงพันธุกรรม มาประยุกต์ใช้ในการหาคำตอบของปัญหาการจัดสมดุลของ

สายงานการประกอบแบบหลายวัตถุประสงค์ ซึ่งหมายถึงการหา รูปแบบของการจัดงานให้กับแต่ละสถานีทำงานเพื่อตอบสนองวัตถุประสงค์ 3 ประการพร้อมๆ กัน คือเพื่อให้สายงานการประกอบมีจำนวนสถานีการทำงานน้อยที่สุด มีรอบเวลาการผลิตน้อยที่สุด และมีความแปรปรวนของภาระงานในแต่ละสถานีการทำงานน้อยที่สุด

จากงานวิจัยดังกล่าวจะเห็นได้ว่าการนำขั้นตอนวิธีเชิงพันธุกรรมไปประยุกต์ใช้ในงานต่างๆ กันอย่างแพร่หลายเพื่อแก้ปัญหาหรือค้นหาคำตอบที่เหมาะสมต่อไป



ภาพที่ 2: ภาพรวมของระบบ

3. วิธีการดำเนินการวิจัย

งานวิจัยการประยุกต์ขั้นตอนวิธีเชิงพันธุกรรมในการกำหนดกำลังทางเรือมีการออกแบบและดำเนินการวิจัยดังต่อไปนี้

3.1 ภาพรวมของระบบ

จากภาพที่ 2 แสดงภาพรวมของระบบ ซึ่งประกอบไปด้วยข้อมูล 2 ส่วนที่จะต้องป้อนให้แก่ระบบมีดังนี้ 1.) ข้อมูลเป้าหมายที่ต้องการจากกองกำลังทางเรือ แสดงดังตารางที่ 1 เป็นตัวอย่างข้อมูล และ 2.) ข้อมูลขีดความสามารถของเรือในแต่ละประเภท แสดงดังตารางที่ 2 เป็นตัวอย่างข้อมูล ซึ่งข้อมูลขีดความสามารถของเรือในแต่ละประเภทนี้จะมีวิธีการประเมินคะแนนดังตารางที่ 3 เป็นตัวอย่างขีดความสามารถ 4 ด้าน ซึ่งในงานวิจัยนี้จะมีด้วยกันทั้งหมด 15 ด้าน [9] และเรือแต่ละประเภทจะมีระดับขีดความสามารถที่แตกต่างกัน โดยจะประเมินคะแนนโดยเสนอวิธีการทหารเรือผู้เชี่ยวชาญด้านเรือรบไทย และข้อมูลทั้ง 2 นี้จะถูกนำไปคำนวณด้วยขั้นตอนวิธีเชิงพันธุกรรมเพื่อค้นหากองกำลังทางเรือที่เหมาะสม ซึ่งระบบจะให้ค้นหาคำตอบ โดยจะให้ผลลัพธ์เป็นจำนวนเรือในแต่ละประเภท

ตารางที่ 1: ตัวอย่างขีดความสามารถที่ต้องการจากกองกำลังทางเรือ

การสนับสนุนอากาศยาน	การเข้าถึงพื้นที่ด้วยเรือยกพล	การค้นหาและช่วยเหลือ	การเคลื่อนย้ายประชาชน
5	5	4	3

ตารางที่ 2: ตัวอย่างขีดความสามารถของเรือในแต่ละประเภท

ประเภทของเรือ	การสนับสนุนอากาศยาน	การเข้าถึงพื้นที่ด้วยเรือยกพล	การค้นหาและช่วยเหลือ	การเคลื่อนย้ายประชาชน
ประเภท 1	0	1	0	2
ประเภท 2	2	1	0	1
ประเภท 3	2	0	1	2
ประเภท 4	1	1	1	0
ประเภท 5	0	2	0	1

ตารางที่ 3: ตัวอย่างการประเมินคะแนนขีดความสามารถด้านต่างๆ

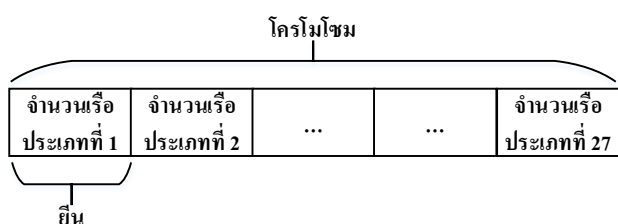
ขีดความสามารถ	คะแนน	คำจำกัดความ
การสนับสนุนอากาศยาน	0	ไม่มีขีดความสามารถในการสนับสนุน
	1	มีคาตฟ้าบิน แต่ไม่มีโรงเก็บอากาศยาน
อากาศยาน	2	มีคาตฟ้าบิน และ โรงเก็บอากาศยาน
การเข้าถึงพื้นที่ด้วยเรือยกพล	0	ไม่มีขีดความสามารถในการยกพล
	1	มีขีดความสามารถในการยกพล
ยกพล	2	มีขีดความสามารถการในบรรทุกเรือรบ
การค้นหาและช่วยเหลือ	0	ไม่มี ฮ. และเรือ RIB ในการค้นหาและช่วยเหลือ
	1	มีเรือ RIB ในการค้นหาและช่วยเหลือ
	2	มี ฮ. และเรือ RIB ในการค้นหาและช่วยเหลือ

การเคลื่อนย้ายประชากร	0	ไม่สามารถเคลื่อนย้ายประชากรได้
	1	เคลื่อนย้ายประชากรไม่เกิน 20 คน/เที่ยว
	2	เคลื่อนย้ายประชากรมากกว่า 20 คน/เที่ยว

3.2 การออกแบบขั้นตอนวิธีเชิงพันธุกรรม

3.2.1 รูปแบบของโครโมโซมและยีน

ในงานวิจัยนี้จะใช้เรือทั้งหมด 27 ประเภท ดังนั้นโครโมโซมนี้จะมียีนทั้งหมด 27 ตัว แสดงดังภาพที่ 3 โดยในแต่ละยีนจะแทนด้วยจำนวนของเรือในแต่ละประเภท ซึ่งรูปแบบของโครโมโซมนี้จะนำไปทำเป็นแม่แบบของกลุ่มประชากรที่ใช้ในการคำนวณต่อไป



ภาพที่ 3: รูปแบบโครโมโซมและยีน

3.2.2 กำหนดประชากรเริ่มต้น (Initial Population)

การกำหนดประชากรเริ่มต้นจะใช้วิธีการสุ่มคำตอบจากจำนวนเรือที่มี โดยจะสุ่มตามจำนวนของประชากรที่ได้กำหนดไว้ (Population Size) ซึ่งการสุ่มนี้จะเป็นต้นแบบเพื่อใช้เป็นจุดเริ่มต้นของกระบวนการขั้นตอนวิธีเชิงพันธุกรรม

3.2.3 การประเมินค่าความเหมาะสม (Fitness Function)

วิธีการประเมินค่าความเหมาะสมเป็นการให้คะแนนสำหรับคำตอบในแต่ละโครโมโซม เพื่อใช้ในการพิจารณาความเหมาะสมของโครโมโซมเพื่อนำไปเข้าสู่กระบวนการขั้นตอนวิธีเชิงพันธุกรรม ซึ่งในงานวิจัยนี้จะใช้สมการดังสมการที่ (1)

$$fitness = \sum_{i=0}^n W_i(A_i - T_i) \quad (1)$$

โดยที่ A_i คือค่าผลรวมขีดความสามารถของเรือด้าน i

T_i คือค่าขีดความสามารถเป้าหมายด้าน i

W_i คือค่าน้ำหนักของขีดความสามารถในแต่ละด้าน i

ซึ่งในงานวิจัยนี้กำหนดให้มค่าน้ำหนักเท่ากันทุกด้าน

และ n เป็นจำนวนขีดความสามารถของเรือ

3.2.4 การคัดเลือก (Selection)

วิธีการคัดเลือกจะใช้วิธีการโดยเก็บคำตอบที่ดีที่สุด P_s (ในที่นี้ใช้ 30%) และอีก 70% จะสุ่มเลือกตามค่าความเหมาะสม (Roulette Wheel Method) ซึ่งการคัดเลือกจะนำมาเป็นโครโมโซมพ่อและแม่ (Parents) เพื่อนำมาใช้ในการสืบสายพันธุ์

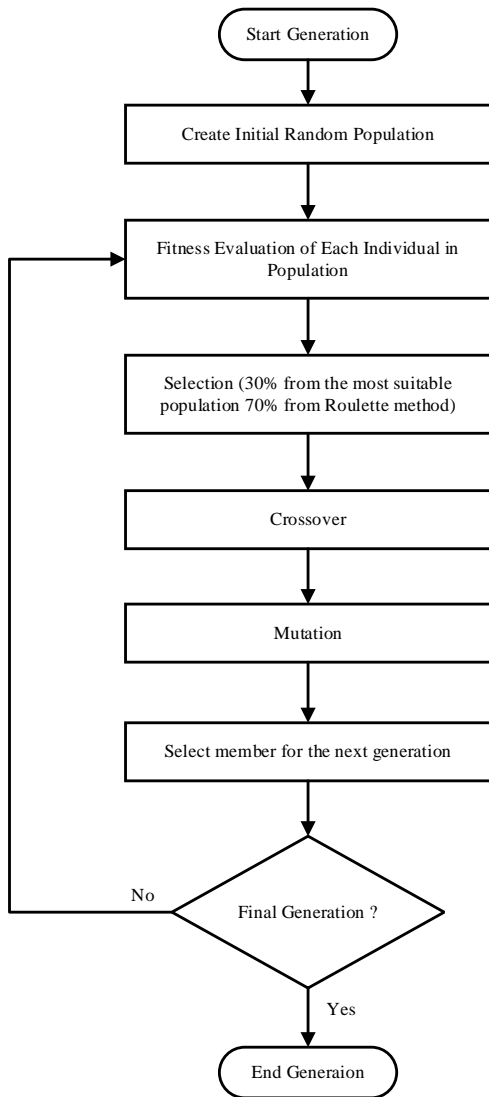
3.2.5 การข้ามสายพันธุ์ (Crossover)

วิธีการข้ามสายพันธุ์เป็นการนำโครโมโซมสองโครโมโซมมาทำการแลกเปลี่ยนยีนกัน จะทำให้ได้โครโมโซมใหม่ขึ้นมาเป็นโครโมโซมลูก ซึ่งจะใช้วิธีการสุ่มในแต่ละคำตอบหรือโครโมโซมว่าจะนำมาใช้ในการผสมหรือไม่ ด้วยความน่าจะเป็น P_c (ในที่นี้ใช้ 20%) เมื่อต้องผสมจะทำการสุ่มคำตอบอื่นที่จะนำมาผสม ซึ่งวิธีการผสมจะใช้การสลับตำแหน่ง เพื่อให้ได้โครโมโซมใหม่ขึ้นมา

3.2.6 การกลายพันธุ์ (Mutation)

วิธีการกลายพันธุ์จะเป็นการเปลี่ยนแปลงบางส่วนของโครโมโซม ซึ่งอาจทำให้คะแนนค่าความเหมาะสม (Fitness Function) ดีขึ้น ซึ่งวิธีการกลายพันธุ์นี้จะใช้วิธีการสุ่มในแต่ละคำตอบหรือโครโมโซมว่าจะนำมาใช้ในการแลกเปลี่ยนยีนหรือไม่ ด้วยความน่าจะเป็น P_m (ในที่นี้ใช้ 30%) และการแลกเปลี่ยนยีนจะใช้วิธีการสุ่มตำแหน่งการสลับยีน

รายละเอียดขั้นตอนต่างๆ สามารถสรุปเป็นกระบวนการทำงานได้ดังภาพที่ 4

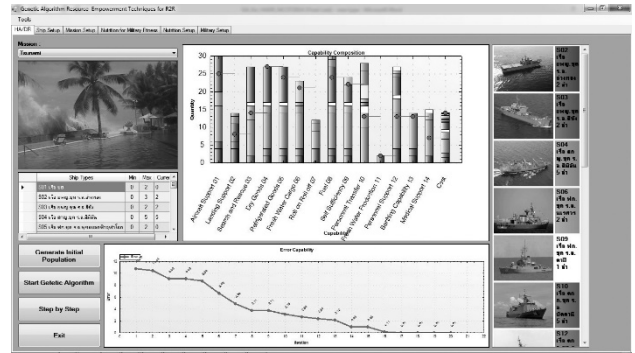


ภาพที่ 4: กระบวนการขั้นตอนวิธีเชิงพันธุกรรม

เริ่มต้นโดยการสร้างประชากรรุ่นแรกโดยวิธีการสุ่มตามจำนวนประชากร ซึ่งจะเป็นคำตอบในตอนแรกเริ่ม และแต่ละโครโมโซมจะถูกพิจารณาค่าความเหมาะสมเพื่อคัดเลือกโครโมโซมไปดำเนินการขั้นตอนวิธีเชิงพันธุกรรมต่อไป และจะได้ประชากรรุ่นใหม่ออกมา ซึ่งประชากรรุ่นใหม่จะถูกกลับมาค้นหาคำตอบตามวัตถุประสงค์ใหม่อีกครั้ง ซึ่งขั้นตอนวิธีดังกล่าวจะกระทำเช่นนี้ไปเรื่อยๆ จนกระทั่งถึงจำนวนรุ่นสุดท้าย หรือตามที่กำหนดไว้

4. ผลการดำเนินการ

จากภาพที่ 5 แสดงตัวอย่างหน้าจอ โปรแกรมกำหนดคองกำลังทางเรือ



ภาพที่ 5: ตัวอย่างหน้าจอโปรแกรมกำหนดคองกำลังทางเรือ

ตารางที่ 4: พารามิเตอร์ที่ใช้สำหรับการทดลอง

พารามิเตอร์	ค่า
จำนวนรุ่น	20
จำนวนประชากร	50
ความน่าจะเป็นในการข้ามสายพันธุ์	0.2
ความน่าจะเป็นในการกลายพันธุ์	0.3



ภาพที่ 6: กราฟแสดงผลค่าผิดพลาด

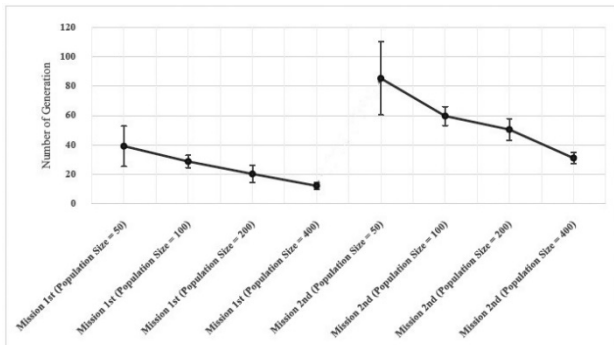
ตารางที่ 5: จำนวนเรือในแต่ละประเภทที่ได้จากการค้นหาคำตอบ

ภารกิจที่	ประเภทและจำนวนเรือ
1	S02=2, S03=2, S04=5, S06=1, S07=1, S09=1, S10=5, S11=1, S12=2, S13=1, S17=1, S18=3, S19=1, S20=1, S21=2, S22=1, S23=3, S24=1
2	S02=3, S03=2, S04=5, S09=1, S10=3, S12=1, S14=2, S15=3, S16=3, S18=2, S19=1, S20=2, S21=1, S22=2, S23=1, S24=1, S25=1

การทดลองงานวิจัยนี้จะใช้ประเภทของเรือทั้งหมด 27 ประเภท โดยที่แต่ละประเภทจะมีขีดความสามารถต่างๆ ทั้งหมด 15 ด้าน โดยจะกำหนดค่าพารามิเตอร์ต่างๆ ของระบบ ดังตารางที่ 4 และทดสอบกับภารกิจ 2 ตัวอย่างที่มีความต้องการขีดความสามารถจากคองกำลังทางเรือที่ต่างกัน จากการทดลองดังภาพที่ 6 พบว่าทั้ง 2 ภารกิจมีค่าความผิดพลาดที่ลดลงเมื่อ

จำนวนรุ่นเพิ่มขึ้น และคำตอบหรือกองกำลังเรือที่ได้ แสดงดังตารางที่ 5 โดยที่ S01, S02, S03,...,S27 แทนด้วยเรือในแต่ละประเภท

จากคำตอบที่ได้จะพบว่ากองกำลังทางเรือมีขีดความสามารถครบถ้วนตามที่ต้องการ นอกจากนี้ยังนำคำตอบไปสอบทานกับเสนาธิการทหารเรือ ซึ่งกองเรือที่ได้นี้สอดคล้องกับกองเรือที่จัดโดยเสนาธิการทหารเรือ



ภาพที่ 7: กราฟแสดงจำนวนประชากรที่มีผลต่อจำนวนรุ่น

จากการทดสอบดังภาพที่ 7 โดยในแต่ละภารกิจจะใช้จำนวนประชากรที่ 50, 100, 200 และ 400 พบว่าเมื่อจำนวนประชากรเพิ่มมากขึ้นมีผลทำให้จำนวนรุ่นที่ใช้ในการค้นหาคำตอบลดน้อยลง ซึ่งในการค้นหาคำตอบจะค้นหาไปจนกว่าคำตอบที่ได้ไม่มีการเปลี่ยนแปลงก็จะหยุดการค้นหาคำตอบ

แม้ว่าจำนวนประชากรต่อรุ่นยิ่งมากจำนวนรุ่นจะน้อยลง แต่เวลาที่ใช้ในการคำนวณต่อรุ่นก็จะมากขึ้น ดังนั้นจำนวนประชากรที่เหมาะสมจึงควรอยู่ระหว่าง 50 ถึง 100

5. สรุป

งานวิจัยนี้เป็นการประยุกต์ใช้ขั้นตอนวิธีเชิงพันธุกรรมมาช่วยในการกำหนดกำลังทางเรือ สำหรับภารกิจการช่วยเหลือด้านมนุษยธรรมและการบรรเทาภัยพิบัติ โดยขั้นตอนวิธีเชิงพันธุกรรมสามารถที่จะค้นหาคำตอบหรือกองกำลังทางเรือที่เหมาะสมได้ และอีกทั้งยังสามารถให้คำตอบได้หลากหลายคำตอบ เพื่อเป็นการประกอบการพิจารณาการตัดสินใจแก่การเลือกชุดกองกำลังทางเรือได้อีกด้วย

ในอนาคตสามารถปรับปรุงขั้นตอนวิธีเชิงพันธุกรรมโดยการกำหนดข้อจำกัดเพิ่มเติม อาทิ เช่น ตำแหน่งที่ตั้งของเรือที่

พร้อมใช้งาน ซึ่งจะถูกนำมาใช้ในการคำนวณด้วย เพื่อเพิ่มข้อพิจารณาในการคัดสรรกองกำลังทางเรือที่เหมาะสมให้มีประสิทธิภาพมากยิ่งขึ้น

นอกจากการประยุกต์ใช้ทางด้านการประกอบกำลังทางเรือแล้ว ทางกองทัพเรือยังมีปัญหาที่สามารถนำระบบนี้ไปแก้ปัญหาอื่นได้ เช่น การคัดเลือกชุดอาหารสำหรับกำลังพลทหารที่ปฏิบัติหน้าที่แตกต่างกัน หรือการจัดวิชาให้ได้ตามหมวดความรู้ที่กำหนดสำหรับอบรมเสนาธิการ

6. กิตติกรรมประกาศ

ผู้วิจัยขอขอบพระคุณ ศูนย์ศึกษายุทธศาสตร์ทหารเรือ กรมยุทธศึกษาทหารเรือ ที่ให้ความอนุเคราะห์เรื่องข้อมูลต่างๆ เกี่ยวกับเรือในกองทัพไทย

เอกสารอ้างอิง

- [1] สำนักงานคณะกรรมการพัฒนาการเศรษฐกิจและสังคมแห่งชาติ. “ภัยพิบัติ : สถานการณ์และแนวโน้ม.” [ออนไลน์]. เข้าถึงได้จาก: <http://www.nesdb.go.th/LinkClick.aspx?fileticket=zMo6ZGfKBsc%3D&tabid=210&mid=735> (วันที่สืบค้นข้อมูล: 5 ธันวาคม 2556)
- [2] วณิช รัตนมณี, “การประยุกต์ใช้เจเนติกอัลกอริทึมในการจัดตารางสอนสำหรับนักศึกษา คณะวิศวกรรมศาสตร์ มหาวิทยาลัยสงขลานครินทร์,” *คณะวิศวกรรมศาสตร์ มหาวิทยาลัยสงขลานครินทร์*, 2554.
- [3] อุดม จันทร์จรัสสุข, “ขั้นตอนวิธีเชิงพันธุกรรมสำหรับปัญหาการจัดเส้นทางรถโดยสาร” *คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง*, 2554.
- [4] Chomtip Pornpanomchai, Verachad Wongsawangtham, Satheanpong Jeungudomporn and Nannaphat Chatsumpun, “Thai Handwritten Character Recognition by Genetic Algorithm (THCRGA)” *International Journal of Engineering and Technology (IACSIT 2011)*, Vol.3, No.2, April 2004.
- [5] กรรณิกา ศิลาพันธ์, “การประยุกต์ใช้เจเนติกอัลกอริทึมในการจัดสมดุลของสายงานการประกอบแบบหลายวัตถุประสงค์” *วิทยานิพนธ์มหาบัณฑิต สาขาวิศวกรรมอุตสาหกรรม บัณฑิตวิทยาลัย, จุฬาลงกรณ์มหาวิทยาลัย*, 2541.
- [6] สืบพงศ์ แสงอุดร และ ชีรเดช วุฒิพรพันธ์, “การประยุกต์ใช้ขั้นตอนวิธีเชิงพันธุกรรมแบบผสมเพื่อแก้ไขปัญหาการจัดตารางการผลิตชนิดสายการประกอบแบบไหลลื่น” *วารสารวิชาการพระ*

จอมเกล้าพระนครเหนือ ปีที่ 22 ฉบับที่ 1 ม.ค.-เม.ย. 2555 หน้า 107-117.

- [7] K. F. Man, K. S. Tang, and S. Kwong, "Genetic Algorithms: Concepts and Applications" *IEEE Trans. On Industrial Electronics*, vol. 43, no. 5, pp. 519-534, 1996.
- [8] Naval Postgraduate School, "An Analysis Of U.S. Navy Humanitarian Assistance And Disaster Relief Operations" *The 8th Annual Acquisition Research Symposium*, 11-12 May, 2011
- [9] Aruna Apte, Keenan D. Yoho, Cullen M. Greenfield and Cameron A. Ingram, "Selecting Maritime Disaster Response Capabilities" *Volume 6, Number 2, July – December*, 2013
- [10] Houqing Lu, Hongjun Zhang Xiaojuan Zhang and Ruixin Han, "An Improved Genetic Algorithm for Target Assignment Optimization of Naval Fleet Air Defense" *Proceedings of the 6th World Congress on Intelligent Control and Automation*, June 21-23, 2006
- [11] Evangelos K. Boulougouris and Apostolos D. Papanikolaou, "Optimisation of the Survivability of Naval Ships by Genetic Algorithms" *3rd Int. EuroConference on Computer Application and Information Technologies in the Maritime Industries COMPIT'04*, 9-12 May, 2004
- [12] James Moffat and Susan Fellows, "Using Genetic Algorithms to Represent Higher Level Planning in Simulation Models of Conflict" *Advances in Artificial Intelligence*, 2011

ระบบสนับสนุนการตัดสินใจสำหรับการเลือกพืชและปุ๋ยอินทรีย์ที่เหมาะสมต่อกัน

Decision Support System for Selecting Crops and Organic Fertilizers Suitable for Each Other

พนารัตน์ ศรีเชษฐา (Phanarut Srichetta)¹ วิชิต ชูอุทธา (Vishit Shuuttha)² และ ธีรภัทร พลศรี (Thiraphat Pholsri)³

สาขาวิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏอุดรธานี

¹phanarut@udru.ac.th, ²mxchiting@gmail.com, ³thiraphat_pholsri@hotmail.com

บทคัดย่อ

ในการเพาะปลูกทั้งพืชสวนและพืชไร่ นั้น เพื่อให้ได้ผลผลิตที่ดีเกษตรกรควรมีความรู้เกี่ยวกับธรรมชาติของพืชนั้น และปุ๋ยที่ใช้ในการบำรุงการเติบโตของพืช แต่เนื่องจากเกษตรกรส่วนใหญ่อยู่ในพื้นที่ชนบทยังขาดความรู้ เจ้าหน้าที่เกษตรเข้าไปดูแลให้คำแนะนำไม่ทั่วถึง ส่งผลทำให้ผลผลิตทางการเกษตรและรายได้ไม่เป็นตามที่คาดหมาย ดังนั้นงานวิจัยนี้มีวัตถุประสงค์เพื่อสร้างระบบสนับสนุนการตัดสินใจสำหรับการเลือกพืชและปุ๋ยอินทรีย์ที่เหมาะสมต่อกัน ซึ่งจะช่วยให้เกษตรกรทราบว่าพืชที่ตนเองปลูกแต่ละประเภท แต่ละช่วงอายุ หรือที่ต้องการบำรุงเฉพาะส่วนนั้น ควรให้ปุ๋ยอินทรีย์ใดบ้างจึงจะเหมาะสม นอกจากนี้ยังช่วยแนะนำว่าปุ๋ยอินทรีย์ที่เกษตรกรมีอยู่นั้นสมควรนำไปใช้กับพืชลักษณะใดบ้าง โดยใช้เทคนิคเพื่อนบ้านที่ใกล้ที่สุดจำนวน k รายการ (k -Nearest Neighbor) ที่อาศัยความแตกต่างระหว่างค่าธาตุอาหารหลักที่พืชต้องการ และค่าธาตุอาหารหลักที่ปุ๋ยอินทรีย์มี โดยระบบงานนี้สามารถช่วยแนะนำพืชและปุ๋ยที่เหมาะสมต่อกันได้ในระดับดี

คำสำคัญ: เทคนิคเพื่อนบ้านที่ใกล้ที่สุด ธาตุอาหาร ปุ๋ยอินทรีย์ พืช ระยะห่างแบบยูคลิดีเซียน

Abstract

To get a good yield of cultivating the garden plants and farm plants, farmers should have knowledge about the nature of plants, and organic fertilizer to enrich the plants' growth. However, since most farmers in rural areas still lack the knowledge and the agriculture officials can't take care of thoroughly, affect the agricultural output and income is not as expected. Therefore, this research aims to build a decision support

system to choose crops and organic fertilizers suitable for each other. This system helps the farmers know that for each species or age of crops they planted or part of crops they want to maintain, which organic fertilizers should be suitable for such crops. It also suggests that existing organic fertilizers the farmers have should apply to which plants using the k -Nearest Neighbor technique based on the distance between the principal nutrients that crops need and the principal nutrients that organic fertilizers have. This system can help to introduce crops and fertilizer appropriate in good level.

Keyword: k -Nearest neighbor technique, Nutrient, Organic fertilizer, Crop, Euclidean distance.

1. บทนำ

ประเทศไทยเป็นประเทศเกษตรกรรมที่หน่วยงานภาครัฐส่งเสริมให้เกษตรกรเพาะปลูกพืชสวนพืชไร่ประเภทต่างๆ ที่เหมาะสมตามสภาพภูมิประเทศและภูมิอากาศเพื่อให้ได้ผลผลิตอย่างสูงสุด ซึ่งพืชแต่ละพันธุ์มีความต้องการธาตุอาหารที่แตกต่างกัน ซึ่งสามารถหาได้จากปุ๋ยที่สามารถเกิดขึ้นเองตามธรรมชาติหรือฝีมือมนุษย์ สำหรับเป็นธาตุอาหารเพื่อบำรุงการเจริญเติบโตแก่พืช หรือเพื่อเน้นการบำรุงส่วนต่างๆ ตามช่วงอายุของพืชประเภทนั้น เช่น การเน้นบำรุงใบ ราก ลำต้น หรือ ดอก เป็นต้น โดยปุ๋ยที่เป็นประโยชน์ทั้งต่อพืชและเกษตรกรคือปุ๋ยอินทรีย์ เนื่องจากเป็นปุ๋ยที่ผลิตจากวัตถุดิบจากธรรมชาติหรืออินทรีย์วัตถุ ทำให้คุณสมบัติของดินดีขึ้นเรื่อยๆ ช่วยลดยื้อธาตุอาหารต่างๆ เอาไว้ไม่ให้สูญหายไปจากดินได้โดยง่าย ส่งเสริมการเจริญเติบโตของจุลินทรีย์ที่มีประโยชน์ในดิน

อย่างไรก็ดี ปุ๋ยอินทรีย์ที่ผลิตจากวัตถุดิบแต่ละประเภทมีธาตุอาหารที่แตกต่างกัน หากเกษตรกรใช้ปุ๋ยอินทรีย์ที่มีธาตุอาหารไม่สอดคล้องกับความต้องการของพืช อาจทำให้ได้ผลผลิตไม่เป็นไปตามที่ตั้งเป้าหมายไว้ ต้องใช้ปุ๋ยมากขึ้นเกินความจำเป็น ดังนั้นสำหรับการทำการเกษตรกรรมเพื่อให้ได้ผลผลิตที่ดีนั้น เกษตรกรควรมีความรู้เกี่ยวกับธรรมชาติของพืชนั้นๆ และปุ๋ยอินทรีย์ที่ใช้กับพืชที่ตนเพาะปลูกว่ามีธาตุอาหารเหมาะสมหรือไม่ แต่เนื่องจากเกษตรกรของประเทศส่วนใหญ่อยู่ในพื้นที่ชนบท เจ้าหน้าที่เกษตรเข้าไปดูแลให้คำแนะนำไม่ทั่วถึง ส่งผลทำให้ผลผลิตทางการเกษตรไม่เป็นตามที่คาดหมาย และเกษตรกรต้องเสียค่าใช้จ่ายและเวลาไปอบรมจากศูนย์การเรียนรู้ต่างๆ ทางกรมเกษตร ซึ่งความรู้ที่ได้อาจไม่ตรงกับที่เกษตรกรต้องการเนื่องจากความหลากหลายของพืชที่ปลูกและวัตถุดิบในการผลิตปุ๋ยอินทรีย์ในแต่ละพื้นที่ที่มีต่างกัน

จากปัญหาดังกล่าว ผู้วิจัยจึงมีแนวคิดในการพัฒนาระบบสนับสนุนการตัดสินใจสำหรับการเลือกพืชและปุ๋ยอินทรีย์ที่เหมาะสมกัน ในรูปแบบของเว็บแอปพลิเคชัน โดยใช้เทคนิคเพื่อนบ้านที่ใกล้ที่สุด k รายการ (k -Nearest Neighbor: k -NN) [1]-[3] ในการวิเคราะห์ว่าปุ๋ยสูตรใดบ้างที่เหมาะสมกับพืชที่ระบุ และพืชชนิดใดบ้างที่เหมาะสมต่อปุ๋ยอินทรีย์ที่ระบุ ซึ่งอาศัยการคำนวณความแตกต่างระหว่างค่าธาตุอาหารหลักที่พืชแต่ละชนิดต้องการและธาตุอาหารที่ปุ๋ยอินทรีย์แต่ละสูตรมี และเลือกเพียง k รายการที่มีค่าระยะห่างน้อยที่สุดเป็นคำตอบประกอบการตัดสินใจของเกษตรกรและผู้ใช้ทั่วไปในการเลือกเลือกพืชและปุ๋ยอินทรีย์เพื่อการเกษตรเพาะปลูกต่อไป

2. ปุ๋ยอินทรีย์และพืช และธาตุอาหารที่จำเป็น

2.1 พืช

พืช เป็นสิ่งมีชีวิตที่เคลื่อนที่ไม่ได้ สามารถพบได้ทั้งบนบกและในน้ำ มีคลอโรฟิลล์ช่วยในการสังเคราะห์แสง บางชนิดไม่สังเคราะห์แสงเองแต่เกาะดูดอาหารจากพืชชนิดอื่น โดยทางการเกษตรกรรมนั้น แบ่งเป็นพืชไร่และพืชสวน [6]

- **พืชไร่** เป็นพืชประเภทไม้ล้มลุกและไม้ทนแล้ง ต้องการน้ำน้อย มีอายุการปลูกและการเก็บเกี่ยวไม่นาน และเมื่อให้ผลผลิตแล้วลำต้นก็จะตาย พืชไร่สามารถจำแนกตาม

ประโยชน์ใช้งานได้เป็น พืชอาหารสัตว์ พืชใช้เมล็ด พืชคลุมดิน พืชหัว พืชสมุนไพร พืชเส้นใย พืชน้ำมัน และพืชน้ำตาล

- **พืชสวน** เป็นพืชที่ปลูกในพื้นที่ไม่มากหรือมากก็ได้ แต่ต้องปฏิบัติดูแลรักษาอย่างประณีต ส่วนมากอายุยืน สามารถเก็บเกี่ยวได้เป็นเวลาหลายปีติดต่อกัน สามารถแบ่งออกเป็น พืชผัก ไม้ผล และ ไม้ดอก ไม้ประดับ

2.2 ปุ๋ยอินทรีย์

ปุ๋ยอินทรีย์ เป็นปุ๋ยที่ได้มาจากอินทรีย์สารที่ผลิตขึ้นโดยกรรมวิธีต่างๆ ก่อนที่จะนำไปใช้ประโยชน์กับพืช แบ่งได้ 3 ประเภทคือ ปุ๋ยคอก ปุ๋ยพืชสด และปุ๋ยหมัก [7]

- **ปุ๋ยคอก** ได้จากมูลของสัตว์ผสมกับเศษอาหารต่างๆ เข้าไปด้วย เมื่อนำมาผ่านกระบวนการหมักแล้วจะได้ปุ๋ยคอกที่มีธาตุอาหารสูง สามารถนำไปใช้ในพื้นที่เพาะปลูกได้อย่างดี

- **ปุ๋ยพืชสด** เกิดจากการไถกลบพืชที่เจริญเติบโตเร็วทนแล้งทนต่อโรคแมลงศัตรูพืชดี เก็บสะสมธาตุอาหารได้เร็วและดี หาได้ง่าย เมล็ดราคาถูกและไม่เป็นพิษต่อพืชชนิดอื่นๆ เช่น พืชตระกูลถั่ว พืชตระกูลข้าว หรือพืชตระกูลหญ้า เมื่อพืชนั้นถูกไถกลบแล้วก็จะถูกย่อยสลายโดยจุลินทรีย์ในดินและปลดปล่อยธาตุอาหารพืชออกมา เกิดความอุดมสมบูรณ์แก่ดินในเวลาเร็วและนาน

- **ปุ๋ยหมัก** ได้จากการผสมเศษพืช เศษขยะมูลฝอย อาจมีซากสัตว์และมูลสัตว์รวมอยู่ด้วย โดยอาศัยกรรมวิธีหมักอย่างง่ายๆ และใช้เวลาในระยะหนึ่งจนเปลี่ยนไปจากรูปเดิมอันเป็นผลมาจากการเปลี่ยนแปลงทางชีวเคมีของจุลินทรีย์ หลังจากนั้นก็สามารถนำเอาปุ๋ยหมักที่ได้ไปใช้ในการปรับปรุงบำรุงดิน



พืชไร่ (ข้าว)



พืชสวน (แคงกวา)



ปุ๋ยคอก



ปุ๋ยหมัก

ภาพที่ 1: ตัวอย่างพืชและปุ๋ยอินทรีย์

2.3 ธาตุอาหารที่จำเป็นต่อพืชและมีในปุ๋ยอินทรีย์

ธาตุอาหารที่จำเป็นสำหรับพืชมีอยู่ 16 ธาตุ โดยเป็นธาตุอาหารที่ได้จากน้ำและอากาศ 3 ธาตุ ธาตุหลัก 6 ธาตุ และธาตุอาหารเสริม 7 ธาตุ [8] โดยธาตุไนโตรเจน ฟอสฟอรัส และโพแทสเซียม เป็นธาตุหลักที่พืชมีความสำคัญต่อการเจริญเติบโตและพืชต้องการมากที่สุด

- **ธาตุไนโตรเจน (N)** ช่วยทำให้พืชเติบโตได้เร็วในระยะแรก ช่วยเสริมใบและลำต้นให้มีสีเขียวเข้ม ช่วยเพิ่มปริมาณโปรตีนให้แก่พืชที่ใช้เป็นพืชอาหาร นอกจากนี้ยังช่วยควบคุมการออกดอกออกผลของพืช

- **ธาตุฟอสฟอรัส (P)** ช่วยให้รากดูดน้ำไปทางเซลล์เข้ามาใช้ประโยชน์ได้มากขึ้น ช่วยแก้ผลเสียเนื่องจากพืชอาจได้รับไนโตรเจนมากเกินไป ส่งเสริมการเจริญเติบโตของรากฝอยและรากแขนงในระยะแรกของการเจริญเติบโต ช่วยเร่งให้พืชแก่เร็ว ช่วยในการออกดอก และสร้างเมล็ดของพืช เพิ่มความต้านทานต่อโรคบางชนิด ทำให้ลำต้นของพืชแข็งแรง ไม่ล้มง่าย

- **ธาตุโพแทสเซียม (K)** ส่งเสริมรากดูดน้ำได้ดีขึ้น ช่วยป้องกันผลเสียหายที่อาจเกิดจากการได้รับไนโตรเจนและฟอสฟอรัสมากเกินไป ช่วยเพิ่มคุณภาพของพืชผักและผลไม้ โดยทำให้พืชมีสีส้ม ขนาด ความหวาน และคงทนต่อสภาวะการเปลี่ยนแปลงของดินฟ้าอากาศ ต่อโรคต่างๆ ได้

ตารางที่ 1 ตัวอย่างพืชและปริมาณธาตุอาหารหลักที่พืชต้องการ

รายละเอียดของพืช		ปริมาณความต้องการ		
ชื่อพืช	อายุ : ส่วนที่บำรุง	N	P	K
กะหล่ำดอก	1 เดือน : ดอก	0.3	1.24	2.34
ผักกาดขาว	1 เดือน : ทั้งหมด	1.37	0.62	0.81
คะน้า	1 เดือน : ทั้งหมด	1.85	0.57	0.79

ธาตุอาหารหลักทั้งสามธาตุนี้ ส่วนใหญ่จะมีการระบุเอาไว้ในผลิตภัณฑ์ปุ๋ยสำหรับพืชไร่พืชสวนสูตรต่างๆ ยี่ห้อต่างๆ โดยปุ๋ยอินทรีย์แต่ละสูตรที่ผลิตจากวัตถุดิบที่แตกต่างกันจะมีค่าธาตุอาหารแตกต่างกัน โดยปุ๋ยอินทรีย์ตามมาตรฐานของกรมวิชาการเกษตร พ.ศ. 2551 นั้นควรมีปริมาณธาตุไนโตรเจนไม่น้อยกว่า 1.0% ธาตุฟอสฟอรัสไม่น้อยกว่า 0.5% และธาตุโพแทสเซียมไม่น้อยกว่า 0.5% โดยน้ำหนัก

ตารางที่ 2 ตัวอย่างปุ๋ยอินทรีย์และปริมาณธาตุอาหารที่ได้

รายละเอียดของปุ๋ยอินทรีย์		ปริมาณที่มี		
ชื่อปุ๋ยอินทรีย์	ประเภท	N	P	K
มูลค้างคาว	ปุ๋ยคอก	1.58	14.28	0.6
ปอเทือง	ปุ๋ยพืชสด	1.98	0.3	2.41
ผักตบชวา	ปุ๋ยหมัก	1.55	0.45	4.9

3. เทคนิคเพื่อนบ้านที่ใกล้ที่สุด k รายการ

เทคนิคเพื่อนบ้านที่ใกล้ที่สุด k รายการ (k-Nearest Neighbor) [1]-[3] โดยทั่วไปเป็นวิธีการที่ใช้ในการจำแนกกลุ่ม (Classification) จำนวน k กลุ่มให้กับข้อมูลทดสอบ (Test set) จากชุดข้อมูลฝึกหัด (Training set) ที่ใช้ในการทำเหมืองข้อมูล และปัญญาประดิษฐ์ โดยการพิจารณาคุณลักษณะของข้อมูลที่หากมีคล้ายคลึงกันมากที่สุด หรือพูดอีกนัยหนึ่งคือมีความแตกต่างกันน้อยที่สุด ถือว่าเป็นข้อมูลที่ควรอยู่ในกลุ่มเดียวกัน ซึ่งเทคนิคนี้มีการประยุกต์ใช้ทางการเกษตร ได้แก่ การจำแนกใบไม้ [4] การจำแนกพันธุ์ยางพารา [5] เป็นต้น โดยขั้นตอนการทำงานของเทคนิคเพื่อนบ้านที่ใกล้ที่สุด k รายการสำหรับการศึกษาค้นคว้าครั้งนี้ มีรายละเอียดดังภาพที่ 2

ข้อมูลนำเข้า: ชุดข้อมูลฝึกหัด (*Train*) ข้อมูลทดสอบ (*Test*) และจำนวนข้อมูลที่ต้องการ (*k*)

ข้อมูลออก: ข้อมูลระยะห่างที่น้อยที่สุด *k* รายการแรก (*D*) และตำแหน่งของ *Train k* รายการแรก (*Index*)

กระบวนการ:

- กำหนดให้ *No* คือจำนวนข้อมูลทั้งหมดใน *Train*
- สำหรับ $i=1$ ขณะที่ $i \leq No$ ให้วนซ้ำโดยเพิ่มค่า *i* ทีละ 1
 - คำนวณระยะห่างระหว่าง $Train_i$ กับ *Test*
 - เก็บค่าระยะห่างที่คำนวณได้ลง D_i
- เรียงลำดับค่าระยะห่างใน *D* จากน้อยไปมาก และเก็บตำแหน่งตัวชี้ของ *Train* ลงใน *Index* ตามการเรียงลำดับระยะห่างใน *D*
- ปรับค่าระยะห่างใน *D* ให้อยู่ในรูปแบบบรรทัดฐาน
- คืนค่า *D* และ *Index* จำนวน *k* รายการแรก

ภาพที่ 2: การทำงานของเทคนิคเพื่อนบ้านที่ใกล้ที่สุด k รายการ

ในขั้นของการคำนวณระยะห่างระหว่างข้อมูล 2 ชุดนั้น สามารถคำนวณได้หลายวิธี ได้แก่ โดยการใช้วิธีระยะห่างยูคลิดีเนียน (Euclidean distance) [3] ดังสมการที่ (1)

$$d(x, y) = \sqrt{\sum_{i=1}^N (x_i - y_i)^2} \quad (1)$$

เมื่อ x, y คือ ข้อมูลชุดที่ x และข้อมูลชุดที่ y ตามลำดับ N คือ จำนวนคุณลักษณะที่มีในข้อมูลทั้งสองชุด

ในขั้นของการเรียงลำดับค่าระยะห่างทั้งหมดที่คำนวณได้นั้น มีหลายวิธีในการเรียงลำดับ โดยการศึกษาในครั้งนี้ใช้วิธีการเรียงลำดับอย่างง่ายคือ การเรียงลำดับแบบฟอง (Bubble sort) [9] ที่ใช้วิธีการเปรียบเทียบข้อมูลที่อยู่ติดกันแล้วสลับตำแหน่งเพื่อให้เรียงลำดับ โดยเริ่มเปรียบเทียบจากตัวที่ 1 กับ 2 และตัวที่ 2 กับ 3 ไปเรื่อยๆ จนกว่าจะครบทุกค่า แล้วทำซ้ำจนกว่าจะไม่มีสลับที่ ดังตัวอย่างในตารางที่ 3 ที่เป็นการเรียงตัวเลข 5 ตัว (6, 4, 7, 9, 1) จากน้อยไปมาก (1, 4, 6, 7, 9)

ตารางที่ 3 ตัวอย่างขั้นตอนการเรียงลำดับข้อมูลแบบฟอง

รอบ	ขั้นตอนการเปรียบเทียบคู่ที่ติดกันและการสลับค่า
1	6,4,7,9,1 → 4,6,7,9,1 → 4,6,7,9,1 → 4,6,7,9,1 → 4,6,7,1,9
2	4,6,7,1,9 → 4,6,7,1,9 → 4,6,7,1,9 → 4,6,1,7,9 → 4,6,1,7,9
3	4,6,1,7,9 → 4,6,1,7,9 → 4,1,6,7,9 → 4,1,6,7,9 → 4,1,6,7,9
4	4,1,6,7,9 → 1,4,6,7,9 → 1,4,6,7,9 → 1,4,6,7,9 → 1,4,6,7,9

ในขั้นการทำให้อัตราข้อมูลอยู่ในรูปแบบบรรทัดฐาน (Normalization) นั้น ทำเพื่อให้ทราบวาระห่างที่หาได้มีความห่างมากหรือน้อยเพียงใดเมื่อเทียบกับค่าสูงสุดและค่าต่ำสุดของข้อมูลทั้งหมด [3] โดยปรับข้อมูลระยะห่างให้อยู่ในช่วงเดียวกันคือ [0, 1] ดังสมการที่ (2)

$$D'_i = \frac{D_i - D_{\min}}{D_{\max} - D_{\min}} \quad (2)$$

เมื่อ D'_i คือค่าระยะห่างของข้อมูล i ค่าใหม่

D_i คือค่าระยะห่างของข้อมูล i ค่าเดิม

D_{\max}, D_{\min} คือค่าระยะห่างมากที่สุดและน้อยที่สุดตามลำดับ

4. การดำเนินการศึกษาและพัฒนาระบบ

4.1 การรวบรวมข้อมูล

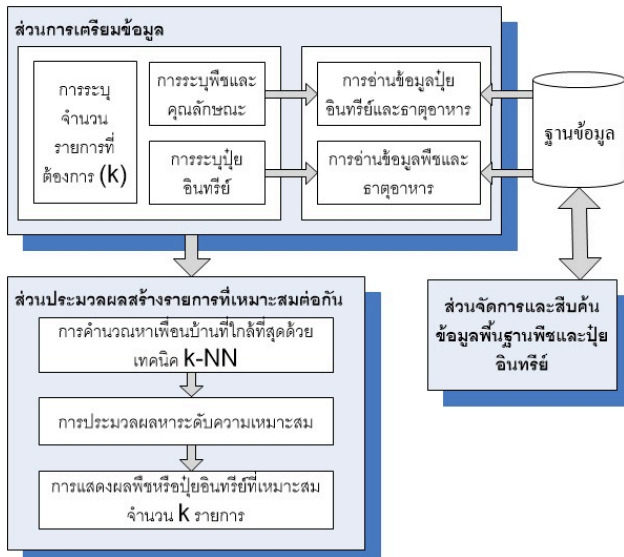
ในการเก็บรวบรวมข้อมูลสำหรับการศึกษาในครั้งนี้ ผู้วิจัยได้ดำเนินการดังนี้ 1) ลงพื้นที่สอบถามเกษตรกรผู้ผลิตปุ๋ยอินทรีย์ทั้งปุ๋ยคอก ปุ๋ยหมัก และปุ๋ยพืชสด และเกษตรกรผู้ปลูกพืชสวนพืชไร่ 2) สอบถามผู้เชี่ยวชาญเกี่ยวกับธาตุอาหารของปุ๋ยอินทรีย์ตามวัตถุดิบที่เลือกใช้ผลิต และธาตุอาหารของพืชที่พืชแต่ละพันธุ์ต้องการ โดยดูทั้งช่วงอายุ หรือส่วนที่ต้องการบำรุงเป็นพิเศษ 3) ศึกษาจากหนังสือและเว็บไซต์ประกอบเพิ่มเติมเกี่ยวกับธาตุอาหาร 4) จัดทำแบบฟอร์มการบันทึกข้อมูลของพืชซึ่งจะมีข้อมูล ชื่อพืช ประเภทพืช อายุพืช ส่วนบำรุง ข้อมูลรายละเอียดเพิ่มเติมของพืชนั้นๆ และค่าธาตุอาหารหลัก (N-P-K) ที่พืชแต่ละชนิดต้องการ โดยเบื้องต้นใช้ข้อมูลตัวอย่างพืช 100 รายการ แบ่งเป็นพืชสวน 50 รายการ และพืชไร่ 50 รายการ และ 5) จัดทำแบบฟอร์มบันทึกข้อมูลของปุ๋ยอินทรีย์ซึ่งจะมีข้อมูล ชื่อปุ๋ยอินทรีย์ ประเภทปุ๋ย และค่าธาตุอาหารหลัก (N-P-K) ที่ได้จากปุ๋ยอินทรีย์แต่ละสูตร โดยใช้ข้อมูลตัวอย่างปุ๋ยอินทรีย์ 30 รายการ แบ่งเป็นปุ๋ยคอก 10 รายการ ปุ๋ยพืชสด 10 รายการ และปุ๋ยหมัก 10 รายการ

4.2 แนวคิดในการพัฒนาระบบงาน

แนวคิดของการพัฒนาระบบสนับสนุนการตัดสินใจการเลือกพืชและปุ๋ยอินทรีย์ที่เหมาะสมต่อกันนั้น จัดทำขึ้นในรูปแบบของเว็บแอปพลิเคชันทำงานผ่านเครือข่ายอินเทอร์เน็ต โดยใช้ภาษา PHP และมี MySQL เป็นระบบการจัดการฐานข้อมูล โดยระบบประกอบด้วยการทำงาน 3 ส่วนหลักแสดงดังภาพที่ 3

ส่วนแรก 1) ส่วนจัดการและสืบค้นข้อมูลพื้นฐานพืชและปุ๋ยอินทรีย์ เป็นส่วนที่ใช้สำหรับการเพิ่ม ลบ แก้ไข และค้นหาข้อมูลพื้นฐานต่างๆ ของพืชและปุ๋ยอินทรีย์ที่จะทำการจัดเก็บไว้ในฐานข้อมูล 2) ส่วนการเตรียมข้อมูล เป็นส่วนการเตรียมค่า k และข้อมูลพืชพร้อมคุณลักษณะเพิ่มเติม หรือข้อมูลปุ๋ยอินทรีย์ตามที่ผู้ใช้ระบุ รวมทั้งอ่านข้อมูลปุ๋ยอินทรีย์ทั้งหมด หรือข้อมูลพืชทั้งหมดจากฐานข้อมูล ก่อนนำไปใช้ในขั้นต่อไป และ 3)

ส่วนประมวลผลสร้างรายการที่เหมาะสมต่อกัน จะใช้ข้อมูลทั้งหมดจากส่วนการเตรียมข้อมูลมาคำนวณหาค่าความแตกต่าง



ภาพที่ 3: องค์ประกอบการทำงานของระบบ

ระหว่างพืชและปุ๋ยอินทรีย์ด้วยเทคนิค k-NN จากนั้นประมวลผลหาระดับความเหมาะสมของค่าที่ได้จาก k-NN แยกเป็น 6 ระดับคือ ดีมาก ดี ปานกลาง พอใช้ น้อย และน้อยมาก ซึ่งเทียบค่าความแตกต่างของระยะห่างที่ผ่านการทำให้อยู่ในรูปแบบบรรทัดฐานที่มีค่าอยู่ในช่วง 0-0.1, 0.11-0.3, 0.31-0.5, 0.51-0.6, 0.61-0.8 และ 0.81-1.0 ตามลำดับ สุดท้ายแสดงผลพืชหรือปุ๋ยอินทรีย์ที่เหมาะสม k รายการเรียงตามระดับความเหมาะสม

4.3 ผลการพัฒนาระบบงาน

ระบบงานที่พัฒนานั้นส่วนติดต่อผู้ใช้จะมีเมนูให้เลือกตามแนวคิด 3 หัวข้อหลัก โดยส่วนการจัดการข้อมูลพื้นฐานของพืชและปุ๋ยอินทรีย์นั้น ผู้ดูแลระบบข้อมูลทางการเกษตรสามารถจัดการเพิ่ม ลบ และแก้ไขรายละเอียดต่างๆ ของพืชและปุ๋ยอินทรีย์ได้ ดังภาพที่ 4 และภาพที่ 5 ตามลำดับ

ภาพที่ 4: หน้าจอส่วนการจัดการข้อมูลพืช

ภาพที่ 5: หน้าจอส่วนการจัดการข้อมูลปุ๋ยอินทรีย์

สำหรับส่วนการเตรียมข้อมูล เป็นส่วนของการเตรียมและกำหนดข้อมูลเกี่ยวกับพืชและข้อมูลเกี่ยวกับปุ๋ยอินทรีย์โดยเกษตรกรหรือผู้ใช้ทั่วไป โดยกรณีต้องการทราบปุ๋ยอินทรีย์ที่เหมาะสมจะนำมาใช้กับพืชที่ตนเองปลูกอยู่ ผู้ใช้สามารถระบุรายละเอียดเกี่ยวกับพืชคือ ชื่อพืช อายุพืช ส่วนที่ต้องการบำรุงเป็นพิเศษ รวมทั้งจำนวนรายการที่ต้องการ เมื่อเลือกให้ระบบทำการวิเคราะห์ ระบบจะไปอ่านข้อมูลปุ๋ยอินทรีย์ทั้งหมดจากฐานข้อมูล และทำงานตามส่วนประมวลผลสร้างรายการที่เหมาะสมต่อกันสำหรับพืชชนิดนั้น ดังตัวอย่างในภาพที่ 6 ที่ได้ปุ๋ยอินทรีย์ 5 รายการตามที่ระบุ โดยแจ้งระดับความเหมาะสมของปุ๋ยนั้นประกอบการตัดสินใจของเกษตรกรต่อไป

พืชที่ท่านเลือกคือ **จิง** อายุ **0** สปีดาร์ ส่วนที่เลือกบำรุงคือ **บำรุงลำต้น**
ลำดับความเหมาะสมที่สุด **5** รายการ

ลำดับ	ชื่อ	ค่าความเหมาะสม	ระดับความเหมาะสม
1	มูลม้า	0.000	ดีมาก
2	ปุ๋ยหมักคั้นมันสำปะหลัง	0.007	ดีมาก
3	ปุ๋ยหมักกกสังคา	0.017	ดีมาก
4	ปุ๋ยหมักมูล ไคโนม	0.017	ดีมาก
5	กกสังคา	0.017	ดีมาก

ภาพที่ 6: หน้าจอส่วนเตรียมข้อมูลพืชและแสดงผลการประมวลผลรายการปุ๋ยอินทรีย์ที่เหมาะสม

ในอีกด้านหนึ่ง หากผู้ใช้ต้องการทราบว่า มีพืชใดบ้างที่เหมาะสมจะใช้ปุ๋ยอินทรีย์สูตรที่ตนมีอยู่ โดยตัวอย่างเตรียมข้อมูลและการแสดงผลรายการพืชที่เหมาะสม แสดงดังภาพที่ 7

วิเคราะห์ข้อมูลจากวัตถุดิบปุ๋ยอินทรีย์

ชื่อวัตถุดิบที่มีอยู่	ปุ๋ยคอกมูลค้างคาว
จำนวนรายการ	5 รายการ
วิเคราะห์ค่าความคล้ายคลึง	

วัตถุดิบที่ท่านเลือกคือ **ปุ๋ยคอกมูลค้างคาว** ได้พืชจำนวน 5 รายการ

ลำดับ	ชื่อ	ค่าความเหมาะสม	ระดับความเหมาะสม
1	ยาสูบ	0.000	ดีมาก
2	กระเจียบ	0.832	น้อยมาก
3	ผักกาดหอม	0.836	น้อยมาก
4	มะเขือเทศ	0.857	น้อยมาก
5	ทุเรียน	0.865	น้อยมาก

ภาพที่ 7: หน้าจอส่วนเตรียมข้อมูลปุ๋ยอินทรีย์และแสดงผลการประมวลผลรายการพืชที่เหมาะสม

ในการประเมินประสิทธิภาพของระบบกระทำโดยผู้เชี่ยวชาญทางการเกษตรพืช 5 ท่านและเกษตรกรและผู้ใช้ทั่วไป 15 ท่าน ประเมินในด้าน 1) ความถูกต้องในการประมวลผลเลือกพืชและปุ๋ยอินทรีย์ที่เหมาะสมต่อกัน 2) ความครอบคลุมปัจจัยทางการเกษตร 3) ความรวดเร็วในการประมวลผล และ 4) ความเหมาะสมในการใช้งาน คะแนนเต็ม 5 ในแต่ละด้าน สำหรับการประเมินโดยผู้เชี่ยวชาญ ได้ค่าเฉลี่ย 4.13 ค่าส่วนเบี่ยงเบนมาตรฐาน 0.61 ถือว่าอยู่ในระดับดี ส่วนการประเมินโดยเกษตรกรและผู้ใช้ทั่วไป ได้ค่าเฉลี่ย 4.43 ค่าส่วนเบี่ยงเบนมาตรฐาน 0.52 อยู่ในระดับดีเช่นกัน

5. บทสรุปและข้อเสนอแนะ

ระบบสนับสนุนการตัดสินใจสำหรับการเลือกพืชและปุ๋ยอินทรีย์ที่เหมาะสมต่อกัน ได้พัฒนาขึ้นเพื่ออำนวยความสะดวกให้แก่เกษตรกรหรือบุคคลทั่วไปที่เพาะปลูกพืช ให้รับทราบว่า มีปุ๋ยอินทรีย์สูตรใดบ้างที่เหมาะสมแก่การนำไปใช้บำรุงพืชที่เกษตรกรปลูกอยู่ตามช่วงอายุและส่วนที่ต้องการบำรุง โดยใช้เทคนิคเพื่อนบ้านที่ใกล้ที่สุด วิเคราะห์ตามค่าธาตุอาหารหลักที่พืชต้องการและธาตุอาหารที่ปุ๋ยอินทรีย์มี ซึ่งจะมีค่าระดับความเหมาะสมของปุ๋ยอินทรีย์ที่ระบบเลือกมาให้เกษตรกรนำไปพิจารณาประกอบการตัดสินใจว่าจะใช้ปุ๋ยอินทรีย์นั้นหรือไม่

หากเกษตรกรไม่สามารถหาปุ๋ยอินทรีย์ในพื้นที่ชุมชนของตนได้ หรือหาได้แต่วัตถุดิบในการผลิตปุ๋ยอินทรีย์สูตรนั้นๆ มีน้อยราคาแพง ก็สามารถพิจารณาปุ๋ยอินทรีย์ในระดับรองลงมาเรื่อยๆ ได้ ในอีกด้านหนึ่ง ระบบช่วยให้เกษตรกรทราบว่า มีพืชชนิดใดบ้างที่เหมาะสมแก่การนำปุ๋ยอินทรีย์ที่เกษตรกรมีอยู่ไปใช้บำรุงบ้าง โดยใช้เทคนิควิธีการแบบเดียวกัน ซึ่งหากพืชชนิดนั้นเกษตรกรไม่ได้ปลูก หรือห้ามเสียดหรือกิ่งพันธุ์มาปลูกไม่ได้ ก็สามารถพิจารณาพืชชนิดอื่นในระดับรองลงมาได้เช่นกัน

อย่างไรก็ดีในการประมวลผลเลือกปุ๋ยอินทรีย์และพืชในการศึกษาครั้งนี้ เป็นการหาผลลัพธ์แบบทั่วไป ยังไม่ได้พิจารณาปัจจัยในด้านสภาพแวดล้อม เช่น สภาพดิน ความชื้นในอากาศ หรือฤดูกาล เป็นต้น มาร่วมในการวิเคราะห์ ซึ่งการพัฒนาในขั้นต่อไปสามารถนำปัจจัยเหล่านี้มาเป็นองค์ประกอบรวม จะทำให้เป็นประโยชน์ต่อการเกษตรพืช เฉพาะเจาะจงลงไปตามลักษณะภูมิประเทศและภูมิอากาศของการเพาะปลูกของเกษตรกรในประเทศมากยิ่งขึ้น

เอกสารอ้างอิง

- [1] B.V. Dasarathy, Nearest Neighbor (NN) Norms: Pattern Classification Techniques, *IEEE Computer Society*, 1991.
- [2] Larose, D.T., *Discovering Knowledge in Data: An Introduction to Data Mining*, John Wiley & Sons, 2005.
- [3] Han, J. and Kamber, M., *Data Mining: Concepts and Techniques*, 2nd ed., Academic Press, 2006.
- [4] J.-X. Du, X.-F. Wang, and G.-J. Zhang, "Leaf Shape Based Plant Species Recognition," *Applied Mathematics and Computation*, vol. 185, 2007.
- [5] ศิริอร อุณเศษ รุ่งทิวา อองคะภาส และพนารัตน์ ศรีเชษฐา, "การระบุพันธุ์ยางพาราด้วยเทคนิคเพื่อนบ้านที่ใกล้ที่สุด" การประชุมทางวิชาการระดับชาติ ด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 8, หน้าที่ 502-509, 2555.
- [6] จิตรภรณ์ รัชพันธุ์, *หลักอนุกรมวิธานพืช*. กรุงเทพฯ : สำนักพิมพ์มหาวิทยาลัยเกษตรศาสตร์. 2548.
- [7] อานัฐ ตันโซ. *เกษตรธรรมชาติประยุกต์: หลักการแนวคิด เทคนิคปฏิบัติในประเทศไทย*. พิมพ์ครั้งที่ 2. ปทุมธานี: ฝ่ายชุมชนและผู้ด้อยโอกาส สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ. 2551.

- [8] สำนักวิจัยและพัฒนาการเกษตรเขตที่ 3 จังหวัดขอนแก่น กรมวิชาการ
เกษตร กระทรวงเกษตรและสหกรณ์. สถิติพื้นที่ปลูกของภาค
ตะวันออกเฉียงเหนือ. 2550.
- [9] ขนิษฐา นามิ. โครงสร้างข้อมูลและอัลกอริทึม. นนทบุรี: ไอดีซีฯ.
2548.

ระบบการจัดการครุภัณฑ์โดยใช้ระบบผู้เชี่ยวชาญแบบอิงกฎเกณฑ์

กรณีศึกษา สำนักวิทยบริการ มหาวิทยาลัยขอนแก่น

Durable Articles Management System with Rule-Based Expert System: A Case Study on Academic Reserves Center of Khon Kaen University

มาตุภูมิ กิตติสุนทรโรภาส(Matupoom Kittisuntaropast)¹, รัชฎา เตชะตานนท์(Tanyada Taychatanont)²,

นิภาวรรณ สาเพชร(Nipawan Sapet)³, นันทน์ภัส เบญจมาศ (Nunnapus Benjamas)⁴

ภาควิชาวิทยาการคอมพิวเตอร์สาขาเทคโนโลยีสารสนเทศและการสื่อสาร

คณะวิทยาศาสตร์มหาวิทยาลัยขอนแก่น

e-mail: ¹Matupoom_k@kkumail.com, ²Tanyadataychatanont@gmail.com

³Nipawan.s@kkumail.com, ⁴Nunnapus@kku.ac.th

บทคัดย่อ

การจัดการครุภัณฑ์เป็นการทำงานที่ย่างยากซับซ้อน และใช้เวลามาก โดยเฉพาะในหน่วยงานใหญ่ที่มีครุภัณฑ์จำนวนมาก ดังนั้นระบบการจัดการครุภัณฑ์ที่มีประสิทธิภาพ เป็นสิ่งจำเป็นอย่างยิ่งที่จะช่วยให้หน่วยงานทำงานได้อย่างมีประสิทธิภาพมากยิ่งขึ้น งานวิจัยนี้ได้วิเคราะห์ ออกแบบ และพัฒนาระบบการจัดการครุภัณฑ์ที่ประกอบด้วย ระบบผู้เชี่ยวชาญสำหรับวินิจฉัยว่าครุภัณฑ์แต่ละรายการควรแทงจำหน่าย หรือ สั่งซื้อหรือไม่โดยการวินิจฉัยครุภัณฑ์ได้ดึงความรู้จากผู้เชี่ยวชาญมาสร้างเป็นฐานความรู้ซึ่งมีโครงสร้างเป็นตารางตัดสินใจและแทนค่าความรู้ในลักษณะของกฎ ผลที่ได้จากการทดสอบและวัดประสิทธิภาพของระบบโดยผู้เชี่ยวชาญจำนวน 11 คนและผู้ใช้ทั่วไปจำนวน 15 คน พบว่า ระบบได้ค่าคะแนนความพึงพอใจเฉลี่ยจากผู้เชี่ยวชาญเท่ากับ 3.18 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.98 และได้ค่าคะแนนเฉลี่ย และส่วนเบี่ยงเบนมาตรฐานจากผู้ใช้ทั่วไปเท่ากับ 3.93 และ 0.80 ตามลำดับ

คำสำคัญ: การจัดการครุภัณฑ์ระบบแบบอิงกฎเกณฑ์
ระบบสนับสนุนการตัดสินใจ

Abstract

The durable articles management system in the organization is inevitably time consuming and effortful, especially the big organization where the amount of durable has been increasing. Thus, the efficient durable articles management system is vital for the productivity of the organizations. The objective of this study was to analyze, design, and develop a durable articles management system which includes a prototype expert system for durable articles disposal and order decision. The knowledge for the expert system was acquired from domain experts, analyzed into decision table structure, and then represented as rules. The system was tested and evaluated by 11 experts and 15 end users. The system achieved an averaged satisfaction score equal to 3.18 with the standard deviation of 0.98 from the experts. The score from the end users is equal to 3.93 with the standard deviation of 0.80

Keywords: Durable Articles Management, Rule-based System, Decision Support System.

1. บทนำ

ปัจจุบันระบบจัดการครุภัณฑ์ มีครุภัณฑ์ประเภทต่างๆ จำนวนมากและการดำเนินการกับข้อมูลของครุภัณฑ์ต่างๆ ที่ยังจัดเก็บอยู่ในรูปแบบของเอกสารที่เป็นกระดาษ ทำให้เกิดข้อผิดพลาดได้ง่าย มีความล่าช้า และเกิดความเสียหายต่อข้อมูลได้ นอกจากนี้การพิจารณาครุภัณฑ์แต่ละรายการว่าควรแทงจำหน่าย หรือ สั่งซื้อหรือไม่ทำโดยผู้เชี่ยวชาญเท่านั้น จากครุภัณฑ์ที่มีหลากหลาย และมีจำนวนมากทำให้การพิจารณาการแทงจำหน่าย และการพิจารณาสั่งซื้อครุภัณฑ์เพิ่มเติม ดำเนินไปได้ช้า และเกิดความผิดพลาดได้ง่าย ดังนั้นงานวิจัยนี้จึงสร้างระบบการจัดการครุภัณฑ์โดยใช้ระบบผู้เชี่ยวชาญแบบอิงกฎเกณฑ์ [1],[2]มาช่วยในการจัดการครุภัณฑ์ทำได้อย่างสะดวก รวดเร็ว และถูกต้องตรงตามความต้องการของผู้ใช้ โดยใช้สำนักวิทยบริการ มหาวิทยาลัยขอนแก่นเป็นกรณีศึกษา

2. ทฤษฎีที่เกี่ยวข้อง

2.1 ระบบสนับสนุนการตัดสินใจ (Decision Support System)เป็นระบบที่จะช่วยผู้บริหารในการตัดสินใจเพื่อกำหนดทิศทางของวัตถุหนึ่งๆซึ่งมีองค์ประกอบที่มีความเกี่ยวข้องกัน[3]

2.2 ครุภัณฑ์(Durable)หมายถึงสินทรัพย์ที่หน่วยงานมีไว้เพื่อใช้ในการดำเนินงานมีลักษณะคงทนทั้งสามารถเคลื่อนที่ได้ และเคลื่อนที่ไม่ได้[4]

2.3 ตารางการตัดสินใจ(Decision table)เป็นเครื่องมือที่ใช้แสดงเงื่อนไขการตัดสินใจและเป็นการเลือกการทำงานหรือกระทำกิจกรรมของเงื่อนไขที่ระบุ[5]

วิธีการตัดสินใจแบบตาราง(Decision Table)

ขั้นตอนการเขียนตารางการตัดสินใจ

1.ระบุเงื่อนไขและกฎสำหรับการตัดสินใจ ตามตารางต่อไป

ตารางที่ 1: ระบุเงื่อนไขและกฎสำหรับตัดสินใจ

เงื่อนไข	กฎสำหรับการตัดสินใจ / การกระทำ
การระบุเงื่อนไขสำหรับการพิจารณาการทำงาน	กฎที่เป็นไปได้ภายใต้เงื่อนไขที่ระบุ
การกระทำที่เป็นไปได้	การระบุการเลือกการกระทำภายใต้กฎเกณฑ์

2. กำหนดจำนวนคอลัมน์ จากผลคูณของกฎการตัดสินใจ
3. นำผลลัพธ์ที่ได้มาหารด้วยค่าที่เป็นไปได้ของเงื่อนไขแรก เพื่อต้องการทราบว่าต้องมีค่าของเงื่อนไขนั้นซ้ำกี่ครั้ง
4. ทำซ้ำขั้นตอนที่ 2 สำหรับเงื่อนไขอื่นๆ แต่ใช้จำนวนเงื่อนไขก่อนหน้าเป็นตัวตั้ง
5. สร้างแถวอน เพื่อเติมค่าการตัดสินใจที่ถูกต้องตามเงื่อนไข
6. ตัดแถวที่ไม่จำเป็นออก เพื่อความสะดวกในการตรวจสอบ

3. วิธีดำเนินการวิจัย

3.1 วิเคราะห์ปัญหาและความต้องการ

วิเคราะห์ปัญหาและความต้องการของผู้ใช้ระบบเพื่อนำมาเป็นแนวทางในการพัฒนาแอปพลิเคชันให้ตรงตามความต้องการของผู้ใช้ระบบโดยศึกษากระบวนการตัดสินใจในการกำหนดทิศทางของครุภัณฑ์ตามที่มีการกำหนดไว้

ศึกษาค้นคว้าทฤษฎีและงานวิจัยต่างๆที่เกี่ยวข้องกับระบบตัดสินใจซึ่งการทำงานมีรูปแบบของกฎเงื่อนไขและวิธีแก้ปัญหาในการเขียนโปรแกรมโดยใช้คำสั่งประเภทSwitch-Caseซึ่งได้รับคำแนะนำจากผู้เชี่ยวชาญด้านการจัดการครุภัณฑ์ เนื่องจากรูปแบบการพิจารณาลักษณะเดียวกันมีการตรวจสอบเงื่อนไขอย่างชัดเจนและได้คำตอบที่ถูกต้องในการกำหนดทิศทางในการแทงจำหน่ายหรือสั่งสินค้าในกรณีที่สินค้าคงคลังต่ำกว่าเกณฑ์ โดยค้นคว้าและศึกษาเทคนิคและเครื่องมือที่ใช้

3.2 การออกแบบบทกวี

3.2.1 สร้างตารางช่วยตัดสินใจจัดทำารแบ่งระบบช่วยตัดสินใจข้อมูลครุภัณฑ์ออกเป็น2ระบบคือการแบ่งจำหน่าย และการสั่งซื้อเพิ่ม

ตารางที่ 2: เงื่อนไขการแบ่งจำหน่าย

เงื่อนไข	ค่าที่เป็นไปได้
1.ความถี่ในการซ่อม	a.น้อยกว่า2ครั้ง b.ระหว่าง2-3ครั้ง c.มากกว่า3ครั้ง
2.ระยะเวลาประกันคงเหลือ	a.เท่ากับ0 b.ระหว่าง1เดือนถึง3 เดือน c.มากกว่า3เดือน
3.ศักยภาพคงเหลือ(หลังหักค่าเสื่อม)	a.เท่ากับ0% b.ระหว่าง1%ถึง20% c.ระหว่าง21%ถึง40% d.มากกว่า40%

ตารางที่ 3: เงื่อนไขการสั่งซื้อเพิ่ม

เงื่อนไข	ค่าที่เป็นไปได้
1.ความถี่ในการเบิก-จ่าย	a.ต่ำ b.ปานกลาง c.สูง
2.สินค้าคงคลัง	a.เท่ากับ0 b.ระหว่าง1หน่วยถึง3หน่วย c.มากกว่า3หน่วย
3.ความถี่ในการซ่อม	a.น้อยกว่า2ครั้ง b.ระหว่าง2-3ครั้ง c.มากกว่า3ครั้ง
4.ระยะเวลาประกันคงเหลือ	a.เท่ากับ0 b.ระหว่าง1เดือนถึง3 เดือน c.มากกว่า3เดือน
5.ศักยภาพคงเหลือ (หลังหักค่าเสื่อม)	a.เท่ากับ0% b.ระหว่าง1%ถึง20% c.ระหว่าง21%ถึง40% d.มากกว่า40%

ระบบที่1 การแบ่งจำหน่ายมีรายละเอียดดังนี้

ขั้นตอนที่ 1 สร้างเงื่อนไขการพิจารณาการแบ่งจำหน่าย โดยสอบถามจากผู้เชี่ยวชาญ ดังตารางที่ 2

ขั้นตอนที่ 2 คำนวณหาจำนวนเงื่อนไขที่เป็นไปได้ทั้งหมด
จำนวนเงื่อนไขที่เป็นไปได้ทั้งหมด=จำนวนค่าที่เป็นไปได้ของจำนวนความถี่ในการซ่อม×ค่าที่เป็นไปได้ของระยะเวลาประกันคงเหลือ ×ศักยภาพคงเหลือ

$$\text{จำนวนเงื่อนไขที่เป็นไปได้ทั้งหมด} = 3 \times 3 \times 4 = 36 \text{ เงื่อนไข}$$

ขั้นตอนที่ 3

- เงื่อนไขที่1ความถี่ในการซ่อม

จำนวนเงื่อนไขทั้งหมดเท่ากับ36มาหารด้วยจำนวนค่าที่เป็นไปได้ของเงื่อนไขที่1คือ3ได้ผลลัพธ์เท่ากับ12

ขั้นตอนที่ 4 (ทำซ้ำ)

-เงื่อนไขที่2 ระยะเวลาประกันคงเหลือ

นำค่าที่เป็นไปได้ของเงื่อนไขนี้คือ3ไปหารจำนวนที่เขียนซ้ำของเงื่อนไขที่2 คือ12ได้ผลลัพธ์เท่ากับ4

-เงื่อนไขที่3ศักยภาพคงเหลือ

นำค่าที่เป็นไปได้ของเงื่อนไขนี้คือ4ไปหารจำนวนที่เขียนซ้ำของเงื่อนไขที่3คือ4ได้ผลลัพธ์เป็น1

ขั้นตอนที่ 5 ตัดเงื่อนไขที่มีผลลัพธ์ซ้ำกันออก จะได้ตารางที่ 4

ระบบที่2 การสั่งซื้อเพิ่ม มีรายละเอียดดังนี้

ขั้นตอนที่ 1 สร้างเงื่อนไขการพิจารณาการแบ่งจำหน่าย โดยสอบถามจากผู้เชี่ยวชาญ ดังตารางที่ 3

ขั้นตอนที่2 คำนวณหาจำนวนเงื่อนไขที่เป็นไปได้ทั้งหมด
จำนวนเงื่อนไขที่เป็นไปได้ทั้งหมด=จำนวนค่าที่เป็นไปได้ของจำนวนความถี่ในการเบิกจ่าย×ค่าที่เป็นไปได้ของสินค้าคงคลัง×ค่าที่เป็นไปได้ของความถี่ในการซ่อม×ค่าที่เป็นไปได้ของระยะเวลาประกันคงเหลือ× ค่าที่เป็นไปได้ของศักยภาพคงเหลือ

$$\text{จำนวนเงื่อนไขที่เป็นไปได้ทั้งหมด} = 3 \times 3 \times 3 \times 3 \times 4 = 324 \text{ เงื่อนไข}$$

ขั้นตอนที่ 3-5 ทำแบบเดียวกันกับระบบที่1

ตารางที่ 4: ผลลัพธ์ที่ได้จากการใช้ระบบคัดสินใจ

	1	2	3	4	5	9	13	17	18	19	20	21	25	29	30	31	32	33
ความถี่	<2	<2	<2	<2	<2	<2	2-3	2-3	2-3	2-3	2-3	2-3	>3	>3	>3	>3	>3	>3
ระยะเวลาประกัน	0	0	0	0	1-3	>3	0	1-3	1-3	1-3	1-3	>3	0	1-3	1-3	1-3	1-3	>3
ศักยภาพ	0	1-20	21-40	>40	-	-	-	0	1-20	21-40	>40	-	-	0	1-20	21-40	>40	-
แท่งจำหน่าย	✓	✓	✓				✓	✓					✓	✓	✓			
เก็บไว้				✓	✓	✓			✓	✓	✓	✓				✓	✓	✓

3.2.2 สร้างกฎ ในโปรแกรมประยุกต์นี้เน้นการออกแบบกฎโดยใช้ระบบแบบอิงกฎเกณฑ์เพื่อคัดเลือกรูทที่ตรงกับเงื่อนไขเพื่อกำหนดทิศทางของครุภัณฑ์พร้อมทั้งระบุโครงสร้างโปรแกรมให้ตรงตามความต้องการของผู้ใช้และมีความยืดหยุ่นในการพัฒนาระบบคัดสินใจต้องพิจารณาเงื่อนไขต่างๆซึ่งทำให้ระบบสามารถวิเคราะห์ข้อมูลโดยมีวิธีการคัดสินใจดังนี้

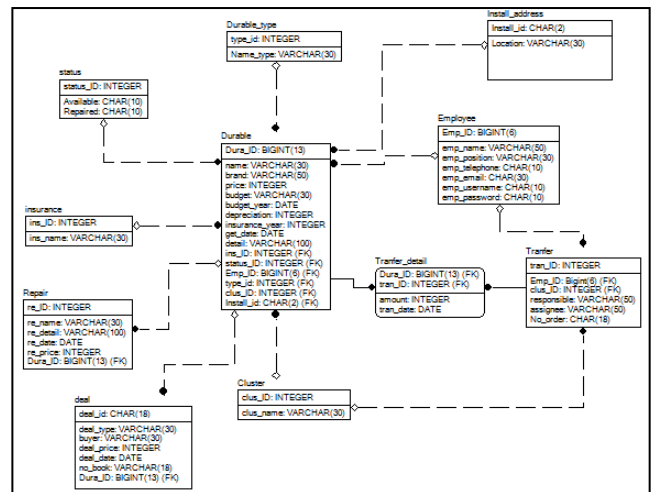
หลักเกณฑ์พื้นฐานที่ใช้คือรูปประโยคSwitch-Caseโดยจะพิจารณาทีละเงื่อนไขของกฎแล้วใช้รูปประโยคIf-Elseเพื่อตรวจสอบเงื่อนไขเพื่อรวมคำตอบจากรูปประโยคSwitch-Case เช่น

```

switch(เงื่อนไข1) {case 2 :คำตอบที่1
case 1 :คำตอบที่ 2
case 0:คำตอบที่3}
switch(เงื่อนไข2) {case 2:คำตอบที่1
case 1:คำตอบที่ 2
case 0:คำตอบที่3}
if(เงื่อนไข1,เงื่อนไข2){ควรจำหน่าย}
    
```

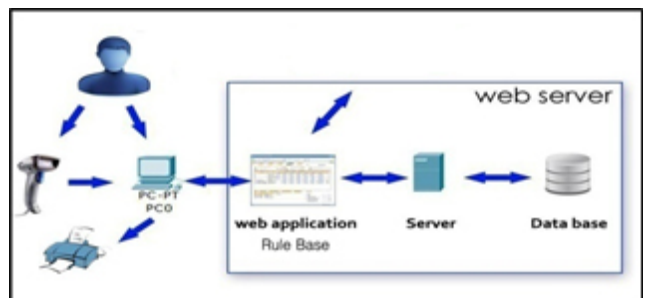
3.3 การวิเคราะห์และออกแบบระบบ

ผู้พัฒนาได้ดำเนินการพัฒนาระบบโดยใช้แผนภาพUMLในการออกแบบผังต่างๆเช่นUse Case Diagram class Diagram[1],[2]ดังภาพที่1



ภาพที่ 1: Class Diagramแสดงความสัมพันธ์ของข้อมูลในระบบ

กระบวนการทำงานของระบบสามารถสรุปเป็นกรอบแนวคิดได้ดังภาพที่2



ภาพที่ 2: กระบวนการทำงานของระบบ

3.4 การพัฒนาระบบ

พัฒนาระบบโดยใช้ระบบปฏิบัติการWindows 7และใช้ภาษาPHPสร้างหน้าเว็บเพจส่วนของโปรแกรมจัดการฐานข้อมูลใช้MySQL Database Serverใช้โปรแกรมAdobe Photoshop

ในการออกแบบตกแต่งหน้าจอกการใช้งานพัฒนาโปรแกรม

ประยุกต์เป็นแบบweb service[6]

3.5 การทดสอบระบบ

เมื่อระบบได้ปรับปรุงแก้ไขเรียบร้อยแล้วขั้นตอนต่อไปคือการทดสอบระบบโดยทดสอบกับบุคคล2กลุ่มคือผู้เชี่ยวชาญ11คน และกลุ่มผู้ใช้งาน15คนซึ่งเก็บเป็นข้อมูลในการประเมินระบบเพื่อตรวจสอบความถูกต้องและความพึงพอใจของผู้ใช้

3.6 การประเมินผลระบบ

3.6.1 การประเมินโดยผู้พัฒนาโปรแกรมโดยใช้วิธีทดสอบแบบแบล็กบ็อกซ์(Black Box Testing)ซึ่งเน้นการทำการของระบบว่าทำงานได้อย่างถูกต้องสมบูรณ์ตามความต้องการของผู้ใช้หรือไม่และเป็นการหาข้อผิดพลาดเพื่อดำเนินการแก้ไขต่อไป

3.6.2 การประเมินความพึงพอใจของโปรแกรมโดยการใช้กฎของไลเคอร์ท(Likert)คือมาตรอันดับ(Rating Scale)เชิงคุณภาพ5ระดับและอันดับเชิงปริมาณ5ระดับซึ่งค่าคะแนนถูกกำหนดตามความเหมาะสม โดยใช้ผู้เชี่ยวชาญและผู้ใช้งานทั่วไปร่วมประเมิน ซึ่งมีรายละเอียดดังตารางที่5

หลังการการประเมินเสร็จได้ทำการวิเคราะห์ทางคณิตศาสตร์โดยใช้หลักการดังต่อไปนี้

3.6.3 ค่าตัวกลางเลขคณิต(Arithmetic Mean)หรือค่าเฉลี่ย(Mean)ดังสูตรที่1

$$\bar{x} = \frac{\sum x}{N}$$

เมื่อ \bar{x} คือค่าเฉลี่ย

$\sum x$ คือผลรวมทั้งหมดของข้อมูล

N คือจำนวนข้อมูลทั้งหมด

3.6.4 ค่าส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) ดังสูตรที่ 2

$$S.D. = \frac{\sqrt{\sum(x-\bar{x})^2}}{N-1}$$

เมื่อ $S.D.$ คือส่วนเบี่ยงเบนมาตรฐาน

X คือค่าของคะแนน

\bar{x} คือค่าเฉลี่ย

N คือจำนวนข้อมูลทั้งหมด

4. ผลการดำเนินงานและผลการประเมิน

ผลการดำเนินงาน

ระบบตัดสินใจประกอบด้วยส่วนสำคัญดังต่อไปนี้

1. ส่วนของการบันทึกจากผู้เชี่ยวชาญเป็นผู้บันทึกเปลี่ยนแปลงแก้ไขกฎตามดุลยพินิจของผู้เชี่ยวชาญ
2. ส่วนของการประมวลผลทิศทางครุภัณฑ์เมื่อผู้ใช้งานทั่วไปกดประมวลผลจะแสดงสถานะว่าครุภัณฑ์ตัวนั้นควรทำอย่างไรเช่นแทงจำหน่ายซื้อเพิ่มเป็นต้น

ผลการประเมินจากการทดสอบระบบจากผู้เชี่ยวชาญ11 คน และผู้ใช้15คน โดยค่าที่ได้มีดังนี้ด้านผู้เชี่ยวชาญมีค่าเฉลี่ยเท่ากับ 3.18(ค่าเบี่ยงเบนมาตรฐานเท่ากับ0.98)และผู้ใช้งานทั่วไปมีค่าเฉลี่ยที่ได้เท่ากับ3.93(ค่าเบี่ยงเบนมาตรฐานเท่ากับ0.80)

ตารางที่ 5: เกณฑ์การให้คะแนนแบบประเมินความพึงพอใจ

ระดับเกณฑ์การให้คะแนน		ความหมาย
เชิงคุณภาพ	เชิงปริมาณ	
สูงมาก	4.51 – 5.00	มีความพึงพอใจระดับดีมาก
สูง	3.51 – 4.50	มีความพึงพอใจระดับดี
ปานกลาง	2.51 – 3.5	มีความพึงพอใจระดับปานกลาง
ต่ำ	1.51 – 2.50	มีความพึงพอใจระดับพอใช้
ต่ำมาก	1.0 – 1.50	มีความพึงพอใจระดับควรปรับปรุง

ตารางที่ 6: สรุปประสิทธิภาพของโปรแกรมจากผู้ใช้งาน

ด้าน	ผู้เชี่ยวชาญ		ผู้ใช้งานทั่วไป	
	X	S.D	X	S.D
1.ความสามารถในการทำงานตามความต้องการผู้ใช้	4.27	0.65	4.33	0.72
2.ด้านหน้าที่ของโปรแกรม	4.18	0.75	4.4	0.63
3.ด้านการใช้งานของโปรแกรม	3.64	0.81	4.53	0.74
4.ด้านประสิทธิภาพของโปรแกรม	3.36	0.67	4.13	0.74
5.ด้านความปลอดภัยของโปรแกรม	3.18	0.98	3.93	0.80

สรุปได้ว่าระบบสนับสนุนการตัดสินใจสำหรับจัดการครุภัณฑ์ โดยใช้กรณีศึกษาจากสำนักวิทยบริการมหาวิทยาลัยขอนแก่นอยู่ในระดับดีเนื่องจากระบบมีความสมบูรณ์ตรงตามความต้องการของผู้ใช้

5. สรุปผลการทำงานของระบบ

ระบบที่พัฒนาขึ้นมาได้มีการตรวจสอบและทดลองใช้งานระบบสามารถใช้งานได้อย่างครบถ้วน และตรงตามความต้องการของผู้ใช้ อีกทั้งยังสามารถช่วยในการตัดสินใจได้อย่างมีประสิทธิภาพโดยสรุปได้ว่าระบบช่วยตัดสินใจของสำนักวิทยบริการมหาวิทยาลัยขอนแก่นมีความพึงพอใจอยู่ในระดับดีและสามารถนำไปต่อยอดในส่วนของการคำนวณกฎ เพื่อให้เกิดความยืดหยุ่นมากขึ้นได้

เอกสารอ้างอิง

- [1] S. Albright, "VBA for Modelers: Developing Decision Support Systems 4ED ," South-Western CENGAGE Learning ,2012.
- [2] James P. Ignizio, *Introduction to Expert Systems The Development And Implementation of Rule-Based Expert Systems*,Mcgraw-hill International Enterprises,1991.
- [3] Chattrakul Sombattheera, "Management Information System for Administration," Faculty of Informatics Mahasarakham University ,Library,2008.
- [4] สุธี สิงห์เสน่ห์. "คู่มือการปฏิบัติงาน การตรวจสอบ และการควบคุมวัสดุ ครุภัณฑ์," หน่วยตรวจสอบภายในสำนักงานปลัดกระทรวงมหาดไทย,2555.
- [5] รุจิรา ธรรมสมบัติ* และทองพูล หีบไธสง, "ระบบสนับสนุนการตัดสินใจเพื่อช่วยค้นหาโรงแรมในอำเภอหัวหิน โดยใช้ตารางการตัดสินใจ," มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, ปีที่พิมพ์ ค.ศ.2009
- [6] พงศ์กรจันทราช. "เอกสารประกอบการสอนรายวิชาการฐานข้อมูล : ภาควิชาเทคโนโลยีสารสนเทศมหาวิทยาลัยฟาร์อีสเทอร์น," (2550). [ออนไลน์]. จาก <http://www.pongkorn.net/dbms>
- [7] พนิดา พานิชกุล และ ณัฐพงษ์ วารีประเสริฐ. "การออกแบบ พัฒนา และการดูแลระบบฐานข้อมูล Database systems - design, development and management," เกทีพี คอมพิวเตอร์ คอนซัลท์, บจก. , ปีที่พิมพ์ ค.ศ.2009
- [8] Doug Davis, Simeon Simeonov and et al., Steve Graham, *Building Web Services with Java*,Developer's Library,2005.

ระบบสนับสนุนการตัดสินใจเกี่ยวกับการส่งออก-นำเข้า ของไทย

Decision support systems for the Export-Import of Thailand

สมเจตน์ ศรีนคร (Somjate Srinakorn)¹ และมาลีรัตน์ โสदानิล (Maleerat Sodanil)²

^{1,2}ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
somjates@moc.go.th¹,msm@kmutnb.ac.th²

บทคัดย่อ

งานวิจัยฉบับนี้มีวัตถุประสงค์เพื่อพัฒนาระบบสนับสนุนการตัดสินใจเกี่ยวกับการส่งออก-นำเข้า สำหรับการวิเคราะห์การส่งออก-นำเข้าของไทย โดยนำข้อมูลจากกรมศุลกากร ซึ่งเป็นข้อมูลรายใบขนสินค้า (Transaction) ที่มีจำนวนมากมาประมวลเพื่อจัดรูปแบบข้อมูลให้เหมาะสม แล้วโหลดเข้าระบบคลังข้อมูล (Data Warehouse) โดยการนำเทคนิค OLAP มาช่วยในการนำเสนอข้อมูลซึ่งจะแสดงในรูปแบบของคิวบ์ (Cube) และนำเทคนิคการทำเหมืองข้อมูล (Data Mining) มาใช้ในการพยากรณ์ปริมาณและมูลค่าการส่งออก-นำเข้า ในการแสดงรายงาน ผู้ใช้งานสามารถดูรายงานได้ตามมุมมองที่ต้องการ นอกจากนี้ข้อมูลยังสามารถแสดงผลในรูปแบบ Dashboard แผนภูมิ แผนที่ ในลักษณะภาพเสมือน (virtual) ได้ด้วย การพัฒนาระบบเป็นลักษณะเว็บแอปพลิเคชัน (Web - Application) โดยแบ่งการทำงาน 2 ส่วน คือ ผู้ใช้งาน สามารถดูรายงานการส่งออก-นำเข้าของไทย และส่วนของผู้ดูแลระบบ ซึ่งจะทำหน้าที่บริหารจัดการข้อมูลผู้ใช้งานและสิทธิ์การเข้าใช้งานระบบ

ผลลัพธ์ที่ได้จาก การพัฒนาระบบนี้ ทำให้ผู้บริหาร หรือเจ้าหน้าที่ภายในกระทรวงพาณิชย์ สามารถที่จะวิเคราะห์สถานะการส่งออก-นำเข้าของไทยได้ง่าย สะดวกรวดเร็ว และถูกต้องมากขึ้น ในส่วนของการพยากรณ์ การพยากรณ์ในกลุ่มสินค้าเกษตรจะมีความแม่นยำกว่าสินค้ากลุ่มอื่น ใช้รูปแบบการพยากรณ์แบบอนุกรมเวลา $Y = T \cdot S \cdot C \cdot I$ สืบเนื่องมาจากสินค้าในหมวดเกษตรจะมีการเปลี่ยนแปลงตามฤดูกาล (Seasonal Variation) ที่ชัดเจนกว่าสินค้าในหมวดอื่น

คำสำคัญ : พิกัดอัตราศุลกากร,คลังข้อมูล,OLAP,อนุกรมเวลา

Abstract

The purpose of this research is to develop the decision support system in respect of export-import for Thai export-import analysis based on a number of the transaction data of The Customs Department used in processing to be prepared in the suitable format and loading them into the data warehouse, using the OLAP technique in assisting in presenting the data as shown in the format of Cube and using the Data Mining technique in forecasting the export-import quantity and value. For presenting the report, the user can see it in line with the points of view as desired. Additionally, the data are also displayed in the form of virtual Dashboard, diagram, and map. The system development is in the form Web Application, dividing into 2 parts, as follows: Part I: Users who can read the export-import report and Part II: System Administrator who will administrate and manage the user data and the right to access the system work.

According to the results of this system development, the administrators or the officers of the Ministry of Commerce can more easily, more conveniently, and more correctly analyze the export-import situation of Thailand. For the forecast, the forecast of the group of the agricultural produces shall be more accurate than other groups due to using the time series forecasting method, as follows : $Y = T \cdot S \cdot C \cdot I$, because the group of the agricultural produces will be seasonally varied as more clearly than other groups of produces.

Keyword : Harmonize,Data warehouse,OLAP,Time - series.

1. บทนำ

ตัวเลขจากธนาคารโลก[1] ระบุว่าในปี 2556 สัดส่วนการส่งออกต่อผลิตภัณฑ์มวลรวมในประเทศ (GDP) ของประเทศไทยมีมูลค่าสูงถึงร้อยละ 75 จากข้อเท็จจริงดังกล่าว จะเห็นได้ว่าไทยมีสัดส่วนการส่งออกต่อGDPในระดับที่ค่อนข้างสูง นอกจากนี้การนำสินค้าเข้ามาเพื่อเป็นปัจจัยในการผลิตสินค้าเพื่อใช้ภายในประเทศและแปรรูปเป็นสินค้าส่งออกไปต่างประเทศต่อไป ทำให้การส่งออกและนำเข้า มีส่วนสำคัญต่อการขยายตัวของเศรษฐกิจของประเทศ

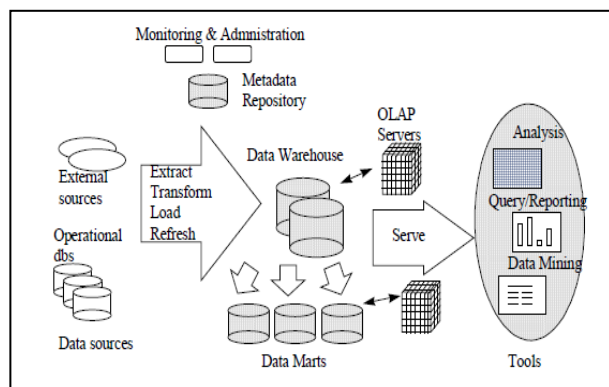
ดังนั้นผู้บริหารและเจ้าหน้าที่ภายในกระทรวงพาณิชย์จำเป็นต้องมีข้อมูลการส่งออก-นำเข้า เพื่อช่วยสนับสนุนการตัดสินใจในการวางแผนงาน และกำหนดนโยบายด้านเศรษฐกิจที่เกี่ยวกับการค้าระหว่างประเทศ ทั้งระยะสั้นและระยะยาว ในปัจจุบันเจ้าหน้าที่ต้องจัดเตรียมข้อมูลที่กระจายอยู่ตามหน่วยงานต่าง ๆ และมีการจัดเก็บที่หลากหลายรูปแบบ เพื่อให้ผู้บริหารใช้สำหรับการตัดสินใจ ซึ่งรายงานเป็นรูปแบบเดิมๆ ที่ใช้เป็นประจำ ทำให้ข้อมูลไม่มีความยืดหยุ่น และเมื่อผู้บริหารต้องการข้อมูลที่แตกต่างออกไป เจ้าหน้าที่จะต้องจัดเตรียมข้อมูลใหม่ ซึ่งจะทำให้เกิดความล่าช้าไม่ตอบสนองต่อความต้องการของผู้บริหารได้ทันเวลา

ด้วยเหตุนี้ผู้จัดทำจึงได้พัฒนาระบบสนับสนุนการตัดสินใจเกี่ยวกับการส่งออก-นำเข้าของไทย สำหรับผู้บริหารและเจ้าหน้าที่ โดยการนำข้อมูลที่เกี่ยวข้องกับการส่งออก-นำเข้า ตั้งแต่อดีตถึงปัจจุบัน มาไว้รวมกันในคลังข้อมูล (Data - Warehouse) ซึ่งระบบ ฯ ที่พัฒนาขึ้นจะใช้งานผ่านระบบเครือข่ายภายในองค์กร (Intranet) จากระบบดังกล่าว ผู้บริหารสามารถดูรายงานได้ ตามมุมมองที่ต้องการ นอกจากนี้ข้อมูลสามารถแสดงผลในรูปแบบของแผนภูมิ หรือภาพเสมือน มีการทำเหมืองข้อมูล (Data Mining) คือ การพยากรณ์ ทำให้ทราบแนวโน้มการส่งออก-นำเข้าของไทย ในอนาคต ซึ่งจะเป็นข้อมูลที่ช่วยให้ผู้บริหารและเจ้าหน้าที่ในกระทรวงพาณิชย์ใช้ในการตัดสินใจ ที่มีประสิทธิภาพ และทันต่อสถานการณ์

2. ทฤษฎีที่เกี่ยวข้อง

2.1 คลังข้อมูล (Data Warehouse)

คือ ที่เก็บข้อมูลขนาดใหญ่ ซึ่งทำการเก็บรวบรวมข้อมูลจากฐานข้อมูลระบบงาน ประจำวัน (Operational Database) และฐานข้อมูลอื่นภายนอกองค์กร (External Database) โดย ข้อมูลที่ถูกจัดเก็บในคลังข้อมูลนั้น มีวัตถุประสงค์ในการนำมาใช้งาน และมีลักษณะของการจัดเก็บแตกต่างไปจากข้อมูลในฐานข้อมูลระบบงานอื่น โดยข้อมูลในคลังข้อมูลจะถูกนำมาใช้เพื่อสนับสนุนการตัดสินใจบริหารงานของผู้บริหาร โดยที่ข้อมูลนั้นอาจมีรูปแบบที่แตกต่างกันแต่มีวัตถุประสงค์ใน การเก็บในทิศทางเดียวกัน [2] สถาปัตยกรรมคลังข้อมูล แสดง ได้ดังภาพที่



ภาพที่ 1: สถาปัตยกรรมคลังข้อมูล

2.2 การประมวลผลในเชิงวิเคราะห์แบบออนไลน์ (Online Analytic Processing)

ประกอบด้วยเครื่องมือที่ช่วยดึง และนำเสนอข้อมูลในหลายมิติ (Multidimensional) ช่วยให้ผู้ให้เห็นข้อมูลแบบ Real Time ได้ในลักษณะต่าง ๆ เช่น ตารางแสดงข้อมูลไขว้ (Pivot Table) หรือตารางแบบเมตริกซ์ (Metric) ตารางรายงานที่ข้อมูลแปรผันได้ตามการป้อนเงื่อนไข (Report Parameter) รายงานที่เจาะลึกได้ (Drill Down) หรือรายงานที่โยงไปยังรายงานย่อยที่มีรายละเอียดอ้างอิงได้ (Drill Through) เป็นรายงานบนหน้าจอที่มีปฏิสัมพันธ์กับผู้ใช้ โดยที่ OLAP ได้รับการออกแบบมาสำหรับผู้ใช้ในระดับของผู้บริหาร หรือหน่วยงานในองค์กร ที่ต้องวิเคราะห์ข้อมูล เพื่อใช้ประกอบการตัดสินใจในระดับสูง [6]

2.3 ลูกบาศก์ (CUBE)

คือ โมเดลข้อมูลของคลังข้อมูล (Data Warehouse) ซึ่งเปรียบเสมือนได้กับรูปลูกบาศก์ ที่มีมุมมองหลากหลาย แต่ละมุมมองทำให้เกิดการคิวรีข้อมูลจาก Data Warehouse ได้หลากหลายแบบคิวรีประกอบด้วยองค์ประกอบที่สำคัญ คือ มุมมองข้อมูล และปริมาณหรือมูลค่าของข้อมูลการผสมผสานของมิติต่าง ๆ ของคิวรี [7]

2.4 Dimensional Modeling

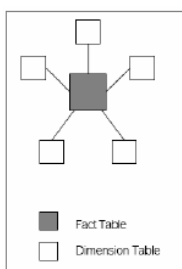
เป็นชื่อเรียกของเทคนิคในการทำให้ฐานข้อมูล ให้ง่ายต่อการทำความเข้าใจโดยการมองภาพของฐานข้อมูลเป็นลูกบาศก์ ที่มี 3,4,5 มิติ หรือมากกว่านั้น ทำให้สามารถจินตนาการ ในการหันหรือแบ่งลูกบาศก์ที่มีลักษณะเหมือนลูกเต๋านี้ได้ นั่นคือสามารถตัดข้อมูลมาวิเคราะห์ดูในช่วงใดก็ได้ และหมุนข้อมูลดูได้จากทุก ๆ ด้านของลูกเต๋านี้ แบ่งได้เป็น 2 แบบ คือ [8]

2.4.1 Star Schema

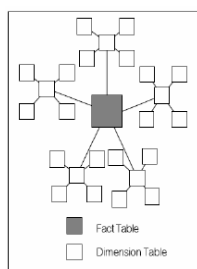
เป็น Dimensional Model รูปร่างคล้ายดาว โดยมี Fact Table ขนาดใหญ่เพียงตารางเดียวอยู่ตรงกลาง และ Dimensional Table เป็นตารางที่มีความสัมพันธ์กับ Fact Table อยู่รอบ ๆ ซึ่งตารางหลักนี้สามารถมีความสัมพันธ์ได้กับอีกหลาย ๆ Dimensional Table แสดงได้ดังภาพที่ 2

2.4.2 Snowflake Schema

มีโครงสร้างที่แตกต่างจากโครงสร้างแบบ Star Schema คือ Dimension table มีหลายระดับ และมีคีย์ที่โยงไปยัง Dimension table อื่นอีก ดังนั้นโครงสร้างแบบนี้จะซับซ้อนมากขึ้น รวมทั้งมีผลให้การสืบค้นยากขึ้นด้วย ลักษณะของข้อมูลจะมีความเป็น Normalized แสดงได้ดังภาพที่ 2 และภาพที่ 3



ภาพที่ 2: Star Schema



ภาพที่ 3: Snowflake Schema

2.5 Dimension Table

เป็นตารางแกนต่าง ๆ ของ Dimensional Database ซึ่งมีลักษณะคล้ายกับตารางประเภท Master ของ OLAP โดย

ภายในประกอบด้วยคอลัมน์ที่ใช้เป็น Key เพื่อเชื่อมโยงไปยัง Fact Table Key และคอลัมน์ที่ให้ความหมายเพิ่มเติมแก่ Entity [2]

2.6 Fact Table

ข้อมูลที่เกี่ยวข้องอยู่ใน Fact Table เรียกว่า Grain Of Fact Table แต่ละเรคคอร์ดใน Fact Table จะเก็บข้อมูลที่วัดได้ของธุรกิจ [8]

2.7 Measured Fact

คือ ข้อมูลที่มีการวัด การประมวลผล หรือได้มีการคำนวณไว้แล้ว จะเก็บอยู่ภายในแต่ละเรคคอร์ดของ Fact Table ปริมาณต่าง ๆ หรือ Measured Fact ที่เพิ่มเข้าไปซึ่งจะมีลักษณะเป็นตัวเลข [5]

2.8 การพยากรณ์รูปแบบอนุกรมเวลาแบบคลาสสิก (Classical Model)

แบ่งอนุกรมเวลาออกเป็น 4 ส่วน ดังนี้

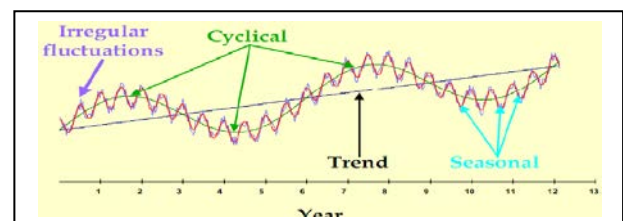
แนวโน้ม (Trend Variation) ใช้สัญลักษณ์ T เป็นการเปลี่ยนแปลงที่เพิ่มขึ้นหรือลดลง

การเปลี่ยนแปลงตามฤดูกาล (Seasonal Variation) ใช้สัญลักษณ์ S เป็นการเปลี่ยนแปลงในช่วงเวลาหนึ่งอาจจะเป็น 1 สัปดาห์ 1 เดือน หรืออื่น ๆ โดยการเปลี่ยนแปลงจะคล้ายกันในช่วงเวลาเดียวกันการเปลี่ยนแปลงตามฤดูกาลจะเห็นได้อย่างเด่นชัดในอนุกรมเวลาของวัตถุดิบ และสินค้าสำเร็จรูป เนื่องจากสิ่งเหล่านี้กระทบกระเทือนได้ง่ายจากสภาวะการณ์ทางธรรมชาติ

การเปลี่ยนแปลงตามวัฏจักร (Cyclical Variation) ใช้สัญลักษณ์ C วัฏจักรทางธุรกิจใอนุกรมเวลาประกอบด้วยช่วงที่แสดงถึงความเจริญและความเสื่อมทางธุรกิจ วัฏจักรอาจเกิดจากเหตุการณ์ภายนอกวงการธุรกิจ เช่น นโยบายของรัฐบาล การเปลี่ยนแปลงทางการเมือง ช่วงของวัฏจักรอาจจะสั้นหรือยาวก็ได้

การเปลี่ยนแปลงที่ผิดปกติ (Irregular Variation) ใช้สัญลักษณ์ I เป็นการเปลี่ยนแปลงที่ไม่ใช่แนวโน้มการเปลี่ยนแปลงตามฤดูกาล และการเปลี่ยนแปลงตามวัฏจักร เช่น การประท้วงทางการเมือง การนัดหยุดงาน สงคราม

ทำการรวมรูปแบบของอนุกรมเวลาในรูปแบบเชิงคูณ (Multiplicative model) $Y = T \cdot S \cdot C \cdot I$ เพื่อนำมาทำการพยากรณ์ [2] แสดงได้ดังภาพที่ 4



ภาพที่ 4: การพยากรณ์รูปแบบอนุกรมเวลา

2.9 ASP.NET

เป็นเทคโนโลยีที่พัฒนามาจาก ASP เพียงแต่ ASP.NET มีการทำเทคโนโลยีของ .NET Framework เข้ามาใช้และเว็บเพจที่พัฒนาขึ้นมาจากเทคโนโลยี ASP.NET จะจัดเก็บอยู่ในไฟล์ .aspx ซึ่งขั้นตอนการทำงาน เมื่อ Web Sever ได้รับความร้องขอการทำงานจาก HTTP โดย WebSever จะนำเอาเฉพาะคำสั่งต่าง ๆ ภายในส่วนของโปรแกรมของ Server Side Script ไปประมวลผลแล้วแปลงผลลัพธ์ที่ได้จากการประมวลผลให้อยู่ในแท็ก HTML ส่งกลับไปพร้อมโปรแกรม Client Side Script เพื่อให้ Brower ประมวลผล อีกครั้ง และ ASP.NET จะมีการแบ่งสถานะการทำงานของแต่ละเว็บเพจ ซึ่งทำให้แยกได้ว่า การเรียกใช้เพจในครั้งนั้น เป็นการเรียกใช้เพจในครั้งแรกหรือไม่ หรือเป็นการเรียกใช้ที่สืบเนื่องมาจากการถูก Submit มารวมทั้งนำเอา ASP.NET Server Control ที่สามารถจดจำค่าที่กำหนดให้กับ Control ที่เกิดขึ้นก่อนหน้านี้ และสามารถตอบสนองต่อเหตุการณ์ต่าง ๆ ที่เกิดขึ้น [4]

3. วิธีการดำเนินงานวิจัย

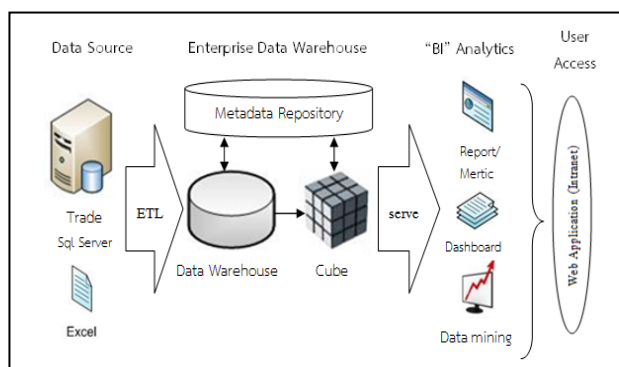
3.1 ศึกษาปัญหาและข้อมูล

ข้อมูลการส่งออก - นำเข้าที่กระทรวงพาณิชย์ได้รับจากกรมศุลกากรเป็นประจำทุกเดือน มาจัดทำรายงานในส่วนของ การวิเคราะห์และแสดงในรูปแบบตารางสถิติ เพื่อใช้ในการแถลงข่าวประจำทุกเดือนและเผยแพร่บนเว็บไซต์ มีจำนวนใบขนสินค้า (Transaction) (มากกว่าสี่ล้านเรคคอร์ด และมีฟิลด์อัตรากรมศุลกากร (Harmonize) ที่ใช้ในการจำแนกประเภทสินค้าตามองค์การกรมศุลกากร โลก มีจำนวนมากกว่าสองหมื่นรายการ ซึ่งข้อมูลดังกล่าวมีปริมาณมากและยังไม่มีมีการจัดกลุ่มหรือจำแนกข้อมูลให้เหมาะตามมุมมองหรือมิติข้อมูลที่ต้องการ

ดังนั้นจึงต้องมีการจัดรูปแบบข้อมูลให้เหมาะสม ด้วยการนำฟิลด์อัตรากรมศุลกากร มาแบ่งกลุ่มสินค้าใหม่ โดยแบ่งสินค้าออกเป็น 2 กลุ่ม คือ ผลิตภัณฑ์ และ ปัจจัยการผลิต ซึ่งมีรายการสินค้ารวมกัน 1,032 รายการ นอกจากนี้ยังมีการแบ่งกลุ่มประเทศ โดยแบ่งกลุ่ม ตามภูมิศาสตร์และกลุ่ม (ทวีป) ความตกลงทางการค้า เพื่อให้ข้อมูลเหมาะสำหรับใช้ในการวิเคราะห์หรือใช้สำหรับการตัดสินใจต่อไป

3.2 การวิเคราะห์และออกแบบระบบงาน

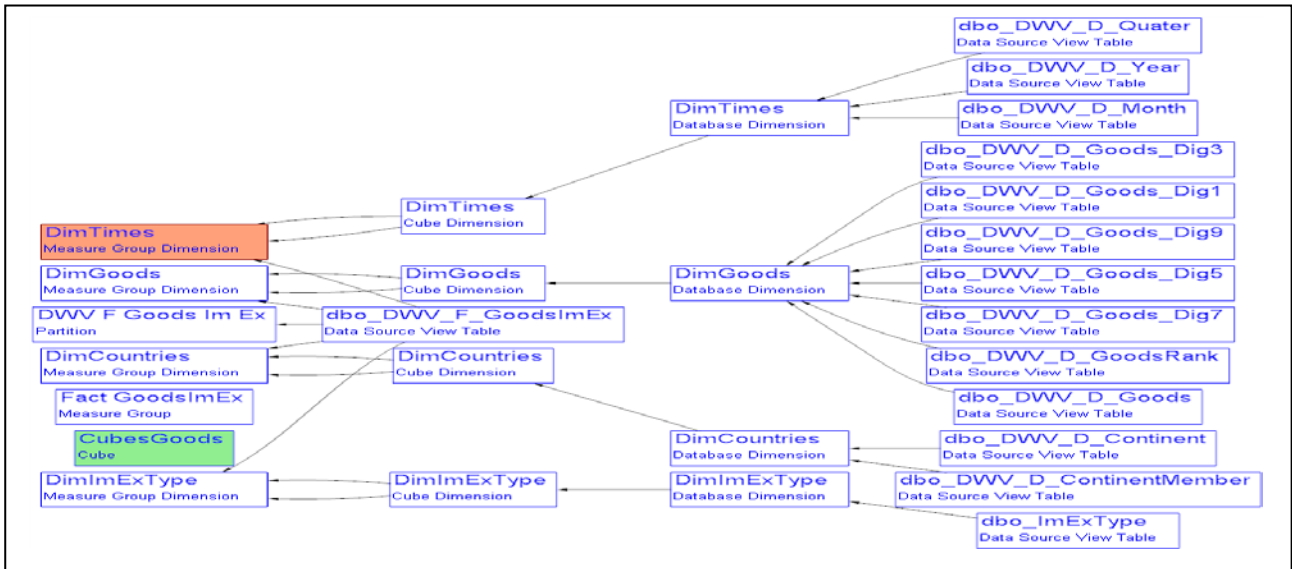
หลักการทำงานในภาพรวมของระบบสนับสนุนการตัดสินใจเกี่ยวกับการส่งออก-นำเข้าของไทย จะรับข้อมูลจากระบบฐานข้อมูล โหลดเข้าคลังข้อมูล (Data Warehouse (โดยนำเทคนิค OLAP มาช่วยในการนำเสนอข้อมูลในลักษณะของคิวบ์ (Cube) ทำให้เห็นมุมมองข้อมูลหลากหลาย มีส่วนของการทำเหมืองข้อมูล (Data Mining) คือการพยากรณ์ปริมาณและมูลค่าการส่งออก-นำเข้า ระบบฯ ที่พัฒนาขึ้นสามารถแสดงรายงานได้หลายรูปแบบ ทั้งแบบ Metric และแบบ Dashboard ทำให้ช่วยสนับสนุนการตัดสินใจของผู้บริหารและเจ้าหน้าที่ภายในกระทรวงพาณิชย์ สามารถที่จะวิเคราะห์สภาวะการส่งออก-นำเข้าของไทยได้ง่าย สะดวกรวดเร็วและถูกต้องมากขึ้น ในการพัฒนาระบบจะเป็นลักษณะเว็บแอปพลิเคชัน (Web Application) (โดยแบ่งการทำงาน 2 ส่วน คือ ผู้ใช้งาน สามารถดูรายงานการส่งออกนำเข้าของไทย- ส่วนผู้ดูแลระบบ ซึ่งจะทำหน้าที่บริหารจัดการข้อมูลผู้ใช้งานและสิทธิ์การเข้าถึงข้อมูล แสดงได้ดังภาพที่ 5



ภาพที่ 5: แนวความคิดในการออกแบบระบบในภาพรวม

3.3 ออกแบบฐานข้อมูล

ผู้พัฒนาได้วิเคราะห์ข้อมูลที่จะดำเนินการประมวลผลและได้ทำการจัดเก็บไว้ในคลังข้อมูลในลักษณะรูปแบบคิวบ์ (Cube) โดยออกแบบ Dimensional Modeling ด้วยโครงสร้างแบบ Snowflake Schema มีการกำหนดตาราง Fact Table พร้อมทั้งเลือกค่าที่จะเป็น Measured Fact และแบ่ง Dimension ออกเป็น 3 มิติ ได้แก่ ประเภทสินค้า, กลุ่มประเทศ และเวลา แสดงในลักษณะเมตาดาต้าของคิวบ์ได้ดังภาพที่ 6



ภาพที่ 6: เมตาดาต้าของคิวบ์

3.4 การพัฒนาระบบ

เป็นลักษณะเว็บแอปพลิเคชัน (Web Application) ด้วยภาษา ASP.NET การเข้าถึงข้อมูลและการใช้งานดำเนินการผ่านระบบเครือข่ายภายในองค์กร

พัฒนาด้วยเครื่องมือ Business Intelligence Development Studio 2008 โดยแบ่งออกเป็น ส่วน ๆ ดังนี้

Server Integration Services (SSIS) สำหรับโหลดข้อมูลจากฐานข้อมูล MS SQL Server 2005 และเอกสารในรูปแบบ Excel เข้าคลังข้อมูล (Data Warehouse)

SQL Server Analysis Services (SSAS) สำหรับสร้าง Cubes และ Dimension ข้อมูลเพื่อใช้สำหรับการวิเคราะห์

SQL Server Reporting Services (SSRS) สำหรับการนำเสนอข้อมูลรูปแบบตาราง และแผนภาพ

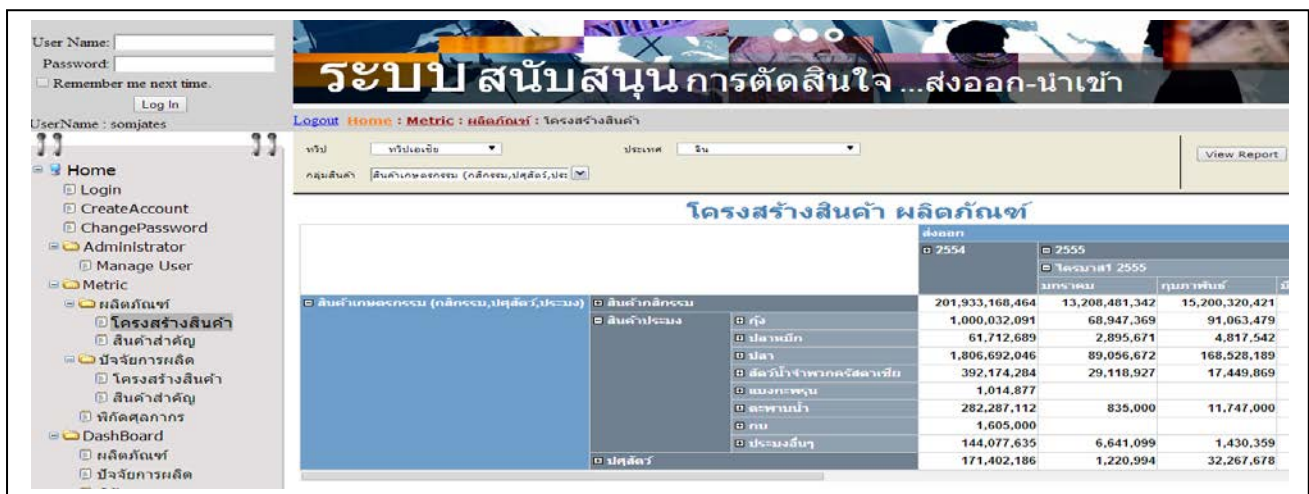
4. ผลลัพธ์การพัฒนาระบบ

4.1 การตรวจสอบสิทธิ์การเข้าใช้ระบบ

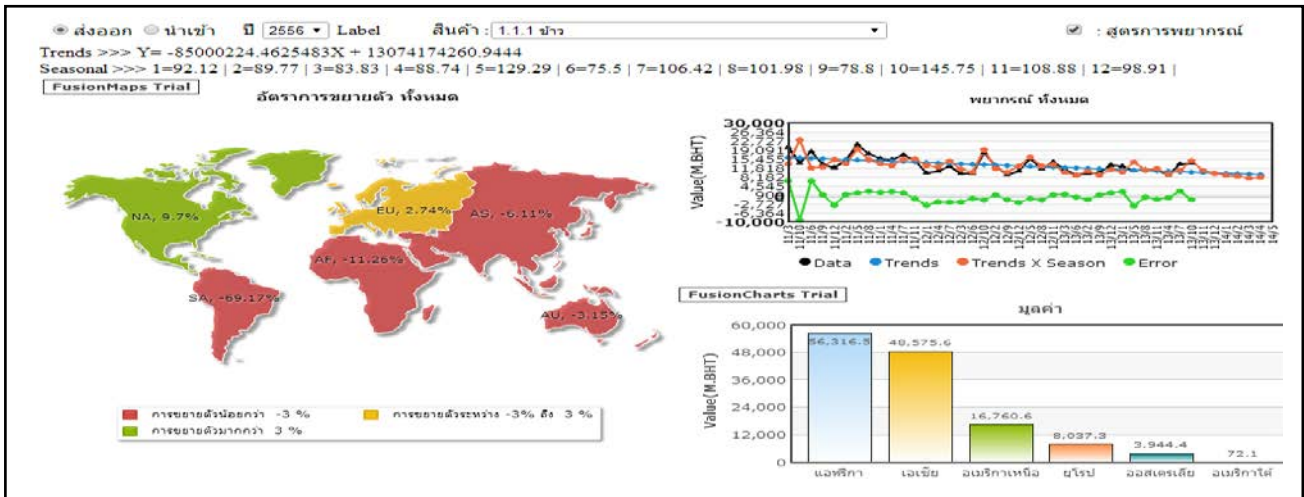
เมื่อเข้าสู่ระบบ จะมีการตรวจสอบสิทธิ์การเข้าใช้งานว่ามีสิทธิ์หรือไม่ และมีระดับผู้ใช้หรือผู้ดูแลระบบ

4.2 รายงานรูปแบบ (Metric)

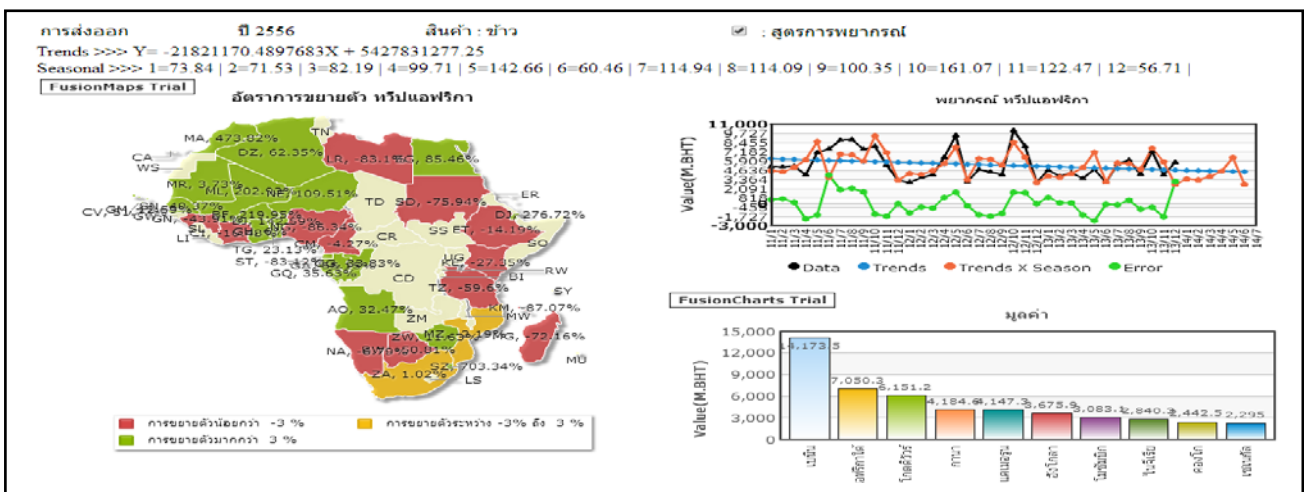
ผู้ใช้งานสามารถดูปริมาณหรือมูลค่าของข้อมูล แบบผสมผสานของมิติต่าง ๆ ของคิวบ์ ตามมุมมองรายงานที่ต้องการได้ เช่น ตัวอย่างรายงานที่ 1 เมื่อผู้ใช้งานสามารถเลือกประเทศและกลุ่มสินค้าที่ต้องการแล้ว รายงานจะแสดงข้อมูลของรายการที่เลือก โดยจะแสดงสินค้าตามแนวโน้ม และสามารถ Drill Down สินค้าเพื่อดูสินค้าย่อย ๆ ได้ ในส่วนแนวตั้งของรายงานจะแสดงปี ซึ่งสามารถ Drill Down ลึกลงไปเพื่อดูข้อมูลเป็นรายไตรมาส และรายเดือนต่อไปได้อีก แสดงได้ดังภาพที่ 7



ภาพที่ 7: ตัวอย่างรายงานรูปแบบ Metric



ภาพที่ 8: ตัวอย่าง Dashboard มูลค่าการส่งออกทวีปต่าง ๆ



ภาพที่ 9: ตัวอย่าง Dashboard มูลค่าการส่งประเทศต่าง ๆ ในทวีปแอฟริกา

4.3 รายงานในรูปแบบ Dashboard

ทำให้วิเคราะห์ข้อมูลได้ง่าย สามารถตัดสินใจได้อย่างรวดเร็ว จากภาพที่ 8 แผนที่แสดงอัตราการขยายตัวของข้าว (สีแดง:ถดถอย ,สีเหลือง:ทรงตัว,สีเขียว:เติบโต) และแสดงมูลค่าการส่งออกข้าวแต่ละทวีปในรูปแบบกราฟแท่ง นอกจากนี้ยังแสดงแนวโน้มการส่งออกในรูปแบบกราฟเส้นของเดือนหน้า ปีหน้า

เมื่อคลิกเลือกทวีปแอฟริกาในแผนที่ กราฟจะแสดงข้อมูลของทวีปแอฟริกา มีมูลค่าการส่งออกข้าว อัตราการขยายตัวของข้าวแต่ละประเทศเป็นเท่าไร และแนวโน้มการส่งออกข้าวไปทวีปแอฟริกา เดือนหน้า ปีหน้า เป็นเท่าไร แสดงได้ดังภาพที่ 9

5. สรุปผล

ระบบที่พัฒนาขึ้น สามารถแสดงรายงานข้อมูลการส่งออก - นำเข้าได้หลากหลายรูปแบบ ผู้ใช้สามารถเลือกรายงานตามมุมมองที่ต้องการ ซึ่งรูปแบบรายงานมีทั้งแบบ Metric ทำให้

ดูข้อมูลแบบผสมผสานของมิติต่างๆของคิวบ์ สามารถ drill down ลงลึกไปในข้อมูลนั้น ๆ ในเชิงลึกได้อีก สามารถอธิบายเป็นรูปธรรมที่ชัดเจน ทำความเข้าใจข้อมูลได้ง่าย ช่วยสนับสนุนในการวิเคราะห์ตัดสินใจได้เร็วขึ้น นอกจากนี้ยังมีส่วนของการพยากรณ์มูลค่าและปริมาณการส่งออก-นำเข้า ทำให้ผู้บริหารรู้แนวโน้มและทิศทางในอนาคต เพื่อใช้เป็นข้อมูลในการกำหนดนโยบายต่อไป

ข้อเสนอแนะ

เนื่องจากการพยากรณ์ ใช้ข้อมูลในอดีตของมูลค่าการส่งออก-นำเข้ามาวิเคราะห์เพียงตัวแปรเดียว ทำให้สินค้าหมวดเกษตรจะมีความแม่นยำกว่าสินค้าหมวดอื่น เนื่องจากมีการเปลี่ยนแปลงตามฤดูกาลที่ชัดเจนกว่าสินค้าในหมวดอื่น ดังนั้นในการพัฒนาครั้งต่อไป ควรพิจารณาตัวแปรอื่น ๆ มาใช้ในการวิเคราะห์ร่วมด้วย เช่น อัตราแลกเปลี่ยนเงินตราต่างประเทศ ราคาน้ำมันโลก เป็นต้น จะได้มีความแม่นยำมากขึ้น

เอกสารอ้างอิง

- [1] The World bank Group, “Exports of goods and services (% of GDP)”, <http://data.worldbank.org/indicator/ne.exp.gnfs.zs>, 2014.
- [2] Surajit Chaudhuri ,Umeshwar Dayal, “An Overview of Data Warehousing and OLAP Technology”,*ACM Sigmod* , pp. 518-525, 1997.
- [3] คงกฤษ ปิ่นทอง, “การพยากรณ์การผลิตชิ้นส่วนยางในรถยนต์”, การค้นคว้าอิสระบริหารธุรกิจมหาบัณฑิต วิชาเอกการจัดการ วิศวกรรมธุรกิจ คณะบริหารธุรกิจ มหาวิทยาลัยรามคำแหง, 2554.
- [4] สุปราณี หือมา, “ระบบคลังข้อมูลการขายคอมพิวเตอร์”, ปัญหาพิเศษวิทยาศาสตร์มหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ ภาควิชาเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2551.
- [5] บุญศรี สว่างประเสริฐ, “ระบบวิเคราะห์และคาดการณ์มูลค่าการค้าโดยเทคนิคโครงข่ายประสาทเทียม”, ปัญหาพิเศษ วิทยาศาสตร์มหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ ภาควิชาเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2554.
- [6] พงษ์ศักดิ์ วงศ์ร่วมฟ้า, “การพัฒนาระบบสนับสนุนการตัดสินใจทางการตลาดของบริษัทวิเวอร์จี จำกัด”, ปัญหาพิเศษวิทยาศาสตร์มหาบัณฑิต สาขาวิชาระบบสารสนเทศเพื่อการจัดการ ภาควิชาการจัดการเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2553.
- [7] สันติ ชัยรัตนธนาภรณ์, “ระบบคลังข้อมูลกองทุนรวมเพื่อช่วยตัดสินใจงานขายโดยใช้เทคนิควิธีการจัดกลุ่ม”, ปัญหาพิเศษ วิทยาศาสตร์มหาบัณฑิต สาขาวิชาระบบสารสนเทศเพื่อการจัดการ ภาควิชาการจัดการเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2554.
- [8] เมธวดี สุจินันท์กุล, “ระบบคลังข้อมูลยอดขายแฟ้มเกจอินเตอร์เน็ตความเร็วสูง”, ปัญหาพิเศษ วิทยาศาสตร์มหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ ภาควิชาเทคโนโลยีสารสนเทศ เทคโนโลยีพระจอมเกล้าพระนครเหนือปีการศึกษา 2551.

ระบบตรวจสอบอาการและบำรุงรักษารถยนต์เบื้องต้น

Basic Automotive Fault Diagnosis and Repair System

ชั้นยัชชก ฐูปเกิด (Tanchanok Toopkird) นฤมล แฟงลีคำ (Narumol Fengseekom)

และ นันท์นภัส เบญจมาศ (Nunnapus Benjamas)

ภาควิชาวิทยาการคอมพิวเตอร์ สาขาเทคโนโลยีสารสนเทศและการสื่อสาร

คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น

tulipaomamm0@gmail.com, aung_idinn2@hotmail.com, nunnapus@kku.ac.th

บทคัดย่อ

เนื่องจากปัจจุบันมีผู้ใช้รถยนต์เป็นจำนวนมาก และผู้ใช้รถยนต์ส่วนใหญ่ไม่มีความรู้เกี่ยวกับการซ่อมรถยนต์เบื้องต้น หรือวิธีบำรุงรักษารถด้วยตัวเอง จึงจำเป็นต้องมีเครื่องมือที่ช่วยให้ผู้ใช้สามารถดูแลรถยนต์ของตัวเองในระดับเบื้องต้นได้ งานวิจัยนี้ได้วิเคราะห์ ออกแบบ และพัฒนาระบบการวินิจฉัยอาการเสียและซ่อมรถยนต์เบื้องต้น ซึ่งประกอบด้วย ระบบผู้เชี่ยวชาญสำหรับการวินิจฉัยปัญหาของรถยนต์ และส่วนการให้ข้อมูลการซ่อมบำรุงรถยนต์เบื้องต้น โดยส่วนการวินิจฉัยปัญหาของรถยนต์ได้ดึงความรู้จากผู้เชี่ยวชาญและคู่มือรถยนต์ มาสร้างเป็นฐานความรู้ซึ่งมีโครงสร้างเป็น Decision tree และแทนค่าความรู้ในลักษณะของกฎ ผลที่ได้จากการทดสอบระบบโดยช่างซ่อมรถยนต์และผู้ขับรถยนต์ พบว่าระบบสามารถใช้งานได้เป็นที่น่าพึงพอใจ

คำสำคัญ: การวินิจฉัยปัญหาของรถยนต์เบื้องต้น ระบบแบบอิงกฎเกณฑ์ ระบบปฏิบัติการแอนดรอยด์

Abstract

Nowadays, the car is very popular vehicles. Most car users do not have the knowledge about basic automotive repair. Therefore, it is necessary to have a tool that allows to take care of the own car. The objective of this study was to analyze, design, and develop a basic automotive fault diagnosis and repair system which includes a prototype expert system for automobile fault diagnosis. The knowledge for the expert system was

acquired from domain experts and automobile manuals, analyzed into decision tree structure, and then represented as rules. The system was tested and evaluated by motor car mechanics and motorists and found to be satisfactory by all users.

Keywords: Automotive Fault Diagnosis, Rule based system, Android Application

1. บทนำ

1.1 บทนำ

เนื่องจากปัจจุบันมีผู้ใช้รถยนต์เป็นจำนวนมาก และผู้ใช้รถยนต์ส่วนใหญ่ไม่มีความรู้เกี่ยวกับการซ่อมรถยนต์เบื้องต้น หรือวิธีบำรุงรักษารถด้วยตัวเอง จึงจำเป็นต้องมีเครื่องมือที่ช่วยให้ผู้ใช้สามารถดูแลรถยนต์ของตัวเอง และสามารถจัดการกับปัญหาในระดับเบื้องต้นที่เกิดขึ้นกับรถยนต์ของตนเองได้ และจากการรวบรวมข้อมูลของกระปุกดอทคอมโดยการวิเคราะห์ข้อมูลของนักวิชาทำให้ทราบว่าในขณะที่เดียวกัน เทคโนโลยีก็ได้เข้ามาเป็นส่วนหนึ่งในบทบาทการดำเนินชีวิตของมนุษย์ โดยเฉพาะเทคโนโลยีในด้านของการสื่อสาร เช่น โทรศัพท์มือถือ เป็นต้น และปัจจุบันโทรศัพท์มือถือ ถือได้ว่าเป็นสิ่งที่มนุษย์พบติดตัวแทบจะตลอดเวลา และโดยมากแล้วมือถือระบบปฏิบัติการแอนดรอยด์ก็เป็นที่ยอมรับกันโดยทั่วไป เพราะใช้งานง่ายและราคาถูก

จากปัญหาต่างๆข้างต้นที่ได้กล่าวมานี้ ผู้พัฒนาจึงได้มีแนวคิดที่จะพัฒนาระบบการวินิจฉัยอาการเสียและซ่อมรถยนต์เบื้องต้น เพื่อให้ผู้ใช้สามารถแก้ไขปัญหาของรถได้ด้วยตัวเอง

และใช้ระยะเวลาในการอยู่กลางท้องถนนให้น้อยลง โดยเลือกพัฒนาแอปพลิเคชันบนระบบปฏิบัติการแอนดรอยด์ เพราะเป็นที่นิยมและใช้กันอย่างแพร่หลาย โดยใช้หลักทฤษฎีระบบแบบอิงกฎเกณฑ์ (Rule-Based System) [1], [2], [3], [4] เข้ามาเป็นเครื่องมือในการพัฒนาแอปพลิเคชัน และใช้กฎต่างๆ ที่สร้างขึ้นจากความรู้ของผู้เชี่ยวชาญหลายๆคน ซึ่งได้มาจากการเก็บข้อมูลสอบถามและสัมภาษณ์ รวมถึงนำความรู้ทางทฤษฎีจากหนังสือมารวบรวมเพื่อสร้างเป็นกฎความสัมพันธ์ของข้อมูล เพื่อนำมาใช้ในการทำนายบอกอาการเสียที่เกิดขึ้น โดยแอปพลิเคชันนี้จะช่วยผู้ใช้ในการวิเคราะห์อาการเสียที่เกิดขึ้นจากบริเวณที่แตกต่างกัน ว่าทั้งหมดที่เกิดอาการเสียขึ้นนั้นมีสาเหตุที่แท้จริงมาจากสิ่งใด

1.2 ขอบเขตของโปรแกรม

- แอปพลิเคชันบนโทรศัพท์มือถือสามารถทำงานได้บนระบบปฏิบัติการแอนดรอยด์ 4.0 ขึ้นไป
- แอปพลิเคชันนี้จะแสดงผลได้คียบโทรศัพท์มือถือหรือแท็บเล็ตที่มีขนาดหน้าจอ 7 นิ้ว
- ในการเลือกตัวเลือกในการตอบคำถามนั้น สามารถเลือกได้เพียงหนึ่งตัวเลือกเท่านั้น
- เนื่องจากในส่วนของค้นหาอยู่บริเวณใกล้เคียงเป็นการแสดงแผนที่ซึ่งต้องดึงข้อมูลแบบออนไลน์ จึงต้องทำการต่อสัญญาณอินเทอร์เน็ต จึงจะสามารถได้
- แสดงแผนที่อยู่ช่อมรถยนต์บริเวณภายในจังหวัดขอนแก่นเท่านั้น
- แอปพลิเคชันนี้ในส่วนของการตรวจเช็คอาการรถสำหรับรถกระบะและรถเก๋งเท่านั้น
- สำหรับวิธีซ่อมแซมและดูแลเบื้องต้นสามารถเลือกอาการรถที่เกิดปัญหาตามที่แอปพลิเคชันระบุไว้เท่านั้น

2. ทฤษฎีที่เกี่ยวข้อง

2.1 ระบบแบบอิงกฎเกณฑ์ (Rule-Based System)

ระบบแบบอิงกฎเกณฑ์เป็นระบบที่ได้รับความนิยมมากในปัจจุบัน ซึ่งเป็นการนำเอาความรู้ของผู้เชี่ยวชาญที่เกิดจากการศึกษาและประสบการณ์เฉพาะด้านของผู้เชี่ยวชาญมาสร้างเป็นกฎในการตัดสินใจทำนายและแก้ไขปัญหาต่างๆ โดยกฎ

ต่างๆที่ได้มานี้จะเป็นเครื่องทำหน้าที่แทนผู้เชี่ยวชาญ แต่ข้อจำกัดคือกฎจะไม่สามารถเปลี่ยนแปลงได้ จึงเหมาะกับการแก้ปัญหาที่ไม่ค่อยเปลี่ยนแปลง ระบบแบบอิงกฎเกณฑ์เป็นเทคนิคที่มีความเข้าใจง่ายและสะดวกในการปรับใช้ โดยในงานวิจัยนี้ผู้วิจัยได้นำทฤษฎีนี้มาใช้ในรูปแบบของเงื่อนไข IF/THEN หรือ อยู่ในรูปอื่นที่เท่าเทียมกัน คือ

IF(ถ้า) <เงื่อนไข> THEN(แล้ว) <ผลที่ตามมา> เช่น

IF สตาร์ตติดตามปกติ

AND รถอืด

AND เปิดไฟติดตามปกติ

AND เครื่องร้อนจัด

AND ควันดำ

AND กินเชื้อเพลิงมากกว่าปกติ

THEN กรองอากาศตัน

อธิบายได้ว่า ถ้านยนต์มีอาการ สตาร์ตติดตามปกติ และรถอืด เปิดไฟหน้าติดตามปกติ เครื่องร้อนจัด ควันดำ และกินเชื้อเพลิงมากกว่าปกติ แสดงว่ามีอาการเสียคือ “กรองอากาศตัน”

2.2 ต้นไม้ตัดสินใจ (Decision Tree)

ต้นไม้ตัดสินใจเป็นเทคนิคที่มีทั้งข้อดีและข้อเสียที่เห็นได้ชัดเจน คือเป็นการกลั่นกรองแบบหยาบโดยการตอบคำถามที่ ใช่/ไม่ใช่ เท่านั้น โดยทางเลือกจะมีสองลักษณะคือ เหมาะสมที่จะนำไปวิเคราะห์ต่อ และไม่เหมาะสมที่จะนำไปวิเคราะห์ต่อในรายละเอียดถัดไป จุดแข็งของต้นไม้ตัดสินใจคือ ไม่จำเป็นต้องพิจารณาเกณฑ์ทั้งหมดว่ามีความสำคัญเท่ากัน จึงช่วยให้ประหยัดเวลาเป็นอย่างมาก โดยการตอบคำถามเพียงไม่กี่คำถาม คำถามจะคล้ายกับเมทริกซ์จัดลำดับความสำคัญ แต่เป็นการตอบว่า ใช่/ไม่ใช่ โดยเป็นการถามคำถามทีละข้อ

3. วิธีดำเนินการวิจัย

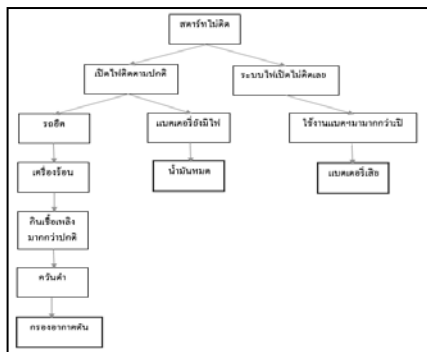
3.1 ศึกษาและค้นคว้า

ศึกษาและค้นคว้าข้อมูล ที่เกี่ยวข้องเกี่ยวกับการทำแอปพลิเคชันบนโทรศัพท์ระบบปฏิบัติการ แอนดรอยด์ วิธีการใช้โทรศัพท์ระบบปฏิบัติการแอนดรอยด์ข้อมูลต่างๆที่เกี่ยวข้องกับรถยนต์ที่จะนำมาใช้ในแอปพลิเคชัน

3.1.1 สอบถามข้อมูลจากผู้เชี่ยวชาญที่มีความรู้และประสบการณ์การทำงานมากกว่า 10 ปี จำนวน 5 คน

โดยสอบถามข้อมูลเกี่ยวกับอาการเสียต่างๆ ของรถยนต์อย่างละเอียด

3.1.2 สร้างแผนภาพต้นไม้ นำข้อมูลที่ได้ทั้งหมดมาสร้างเป็นแผนภาพต้นไม้ เพื่อหาความสัมพันธ์ของข้อมูล ผู้วิจัยได้นำเอาหลักการของต้นไม้ตัดสินใจ (Decision Tree) มาใช้ในการสร้างและจำแนกกฎดังเช่นภาพที่ 1 ที่แสดงส่วนหนึ่งในกิ่งของต้นไม้



ภาพที่ 1 : Decision Tree แสดงกฎความสัมพันธ์ของข้อมูลบางส่วนจากทั้งหมด

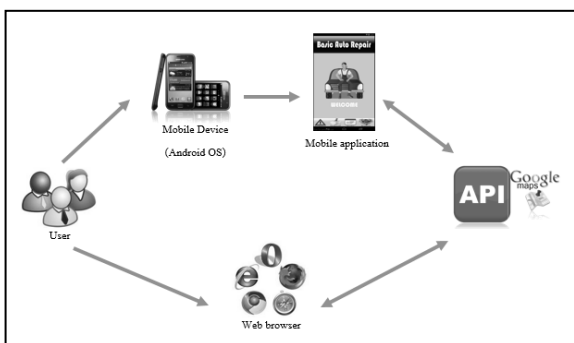
3.1.3 จัดกลุ่ม และสร้างกฎ สร้างกฎของข้อมูลจากแผนภาพต้นไม้ที่สร้างขึ้น

3.2 วางแผนวิเคราะห์ระบบงาน

วางแผนและวิเคราะห์ระบบงานที่จะพัฒนาและออกแบบ Use Case, Use case Scenario, Sequence Diagram, Activities Diagram

3.3 ออกแบบ วิเคราะห์ระบบ

ออกแบบและออกแบบ Prototype Design, หน้า Interface กระบวนการทำงานของระบบสามารถสรุปได้ ดังภาพที่ 2



ภาพที่ 2 : กระบวนการทำงานของระบบ

3.4 พัฒนาระบบ

พัฒนาในส่วนต่างๆของแอปพลิเคชัน และในส่วนของการใช้งาน โดยพัฒนาโดยใช้ระบบปฏิบัติการ Windows 8 และใช้ Android SDK [7] โดยจะยึดโครงสร้างของภาษาจาวาในการพัฒนา และใช้โปรแกรม Adobe Photoshop CS 6 ในออกแบบหน้า ตกแต่งสีสันและภาพต่างๆ ในแอปพลิเคชัน

3.5 ทดสอบและปรับปรุง

ทดสอบความถูกต้องและหาข้อผิดพลาดภายในแอปพลิเคชัน ทดสอบการใช้งานได้จริงของแอปพลิเคชันที่สร้างขึ้น เมื่อพบข้อผิดพลาดก็ทำการปรับปรุงแก้ไขแอปพลิเคชันให้สมบูรณ์

3.5.1 จัดทำแบบสอบถาม ให้ผู้ใช้ทดลองใช้และเก็บข้อมูลลงในแบบสอบถาม โดยแบ่งผู้ใช้ออกเป็นสองกลุ่มคือบุคคลทั่วไป และช่างผู้เชี่ยวชาญ (ช่างซ่อมรถ)

3.5.2 ประเมิน และสรุปผลแบบสอบถามทั้งหมดที่จัดทำขึ้น เพื่อหาข้อผิดพลาดและแก้ไข

3.6 การทดสอบระบบ

ทำการทดสอบระบบเพื่อหาจุดบกพร่อง โดยทดสอบกับตัวอย่าง 2 กลุ่มคือ กลุ่มผู้เชี่ยวชาญ (ช่างซ่อมรถ) จำนวน 5 คน และกลุ่มผู้ใช้งานทั่วไปจำนวน 15 คน โดยใช้แบบประเมินในการเก็บรวบรวมข้อมูล

3.7 การประเมินผลระบบ

3.7.1 ทำการทดสอบระบบเพื่อหาจุดบกพร่อง โดยใช้วิธีทดสอบ แบบกล่องดำ (Black box Testing) กับระบบงานซึ่งเป็นกระบวนการทดสอบระบบโดยภาพรวมทั้งหมด โดยทำการทดสอบการทำงานแต่ละฟังก์ชันการทำงาน รวมทั้งหาข้อบกพร่อง และทำการปรับปรุงแก้ไขให้ดีขึ้น

3.7.2 การประเมินความพึงพอใจ Likert Scale แบ่งความพึงพอใจออกเป็น 5 ระดับ และนำข้อมูลดังกล่าวมาประมวลผลโดยใช้การวิเคราะห์ค่าเฉลี่ย (Mean) และส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) และแปลข้อมูลนำเสนอในรูปแบบของตารางและคำบรรยาย โดยใช้เกณฑ์ในการให้คะแนน และการวิเคราะห์ดังตารางที่ 1 นี้

ตารางที่ 1: เกณฑ์ในการให้คะแนน และการวิเคราะห์

เกณฑ์การให้คะแนน		การแปลความหมาย
ระดับเกณฑ์	ระดับ	
4.51 – 5.00	ดีมาก	แอปพลิเคชันที่พัฒนามีความพึงพอใจในระดับดีมาก
3.51 – 4.50	ดี	แอปพลิเคชันที่พัฒนามีความพึงพอใจในระดับดี
2.51 – 3.50	ปานกลาง	แอปพลิเคชันที่พัฒนามีความพึงพอใจในระดับปานกลาง
1.51 – 2.50	น้อย	แอปพลิเคชันที่พัฒนามีความพึงพอใจในระดับพอใช้
1.00– 1.50	น้อยมาก	แอปพลิเคชันที่พัฒนามีความพึงพอใจในระดับควรปรับปรุง

นำผลที่ได้จากการประเมินมาทำการวิเคราะห์โดยใช้หลักการทางสถิติ การหาค่าเฉลี่ย (Mean) และค่าเบี่ยงเบนมาตรฐาน (Standard Deviation) ใช้ในการหาค่าเฉลี่ยน้ำหนักของความคิดเห็นของกลุ่มตัวอย่าง จากข้อคำถามแบบมาตราส่วนประมาณค่า (Rating Scale)

การหาค่าเฉลี่ยเลขคณิตสำหรับข้อมูล โดยใช้สูตรดังนี้

$$\bar{X} = \frac{\sum x}{n}$$

เมื่อ \bar{X} คือ ค่าเฉลี่ย

$\sum x$ คือ ผลรวมทั้งหมดของข้อมูล

n คือ จำนวนข้อมูลทั้งหมด

ค่าเบี่ยงเบนมาตรฐาน (Standard Deviation)

$$S.D. = \frac{\sqrt{\sum(x-\bar{X})^2}}{N}$$

เมื่อ S.D. คือ ส่วนเบี่ยงเบนมาตรฐาน

X คือ ค่าของข้อมูลแต่ละตัวหรือจุดกึ่งกลางชั้นแต่ละตัว

\bar{X} คือ ค่าเฉลี่ย

N คือ จำนวนข้อมูลทั้งหมด

4. ผลการดำเนินงาน และผลการประเมิน

4.1 ผลการดำเนินงาน

ส่วนประกอบของแอปพลิเคชัน มีดังต่อไปนี้

4.1.1 ส่วนของระบบตรวจสอบอาการรถยนต์ โดยแบ่งแยกเป็น 4 ระบบดังนี้ ระบบเครื่องยนต์ ระบบช่วงล่าง ระบบเบรกและคลัตช์ ระบบไฟฟ้าและแอร์ แสดงดังภาพที่ 3



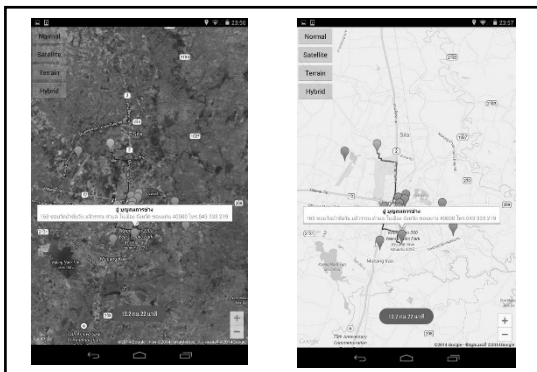
ภาพที่ 3 : ส่วนของระบบตรวจสอบอาการรถ

4.1.2 ตัวอย่างหน้าคำถามซึ่งจะแสดงคำถามต่างๆ เกี่ยวกับอาการของรถยนต์ โดยผู้ใช้งานจะต้องตอบคำถามที่ละเอียดจนกว่าจะครบตามกฎที่ระบบสร้างไว้แสดงดังภาพที่ 4



ภาพที่ 4 : ตัวอย่างเบื้องต้นของอาการระบบเครื่องยนต์

4.1.5 ส่วนของค้นหาอยู่บริเวณในบริเวณใกล้เคียง เป็นส่วนที่ช่วยอำนวยความสะดวกในการค้นหาที่ตั้งของอู่ที่อยู่ในบริเวณใกล้เคียงกับจุดที่ผู้ใช้อยู่ในขณะนั้นมากที่สุด โดยเมื่อผู้ใช้เปิดหน้าขึ้นมา จะพบที่ตั้งของอู่ต่างๆทันที และเมื่อคลิกเลือกที่อู่ใด ก็จะสามารถทราบระยะทาง และรายละเอียดต่างๆ ของอู่ โดยสามารถดูได้ในหลายมุมมอง เช่น มุมมองดาวเทียม และรูปแบบมุมมองแบบแผนที่อีกด้วยแสดงดังภาพที่ 5



ภาพที่ 5 : ตัวอย่างของการค้นหาบริเวณใกล้เคียง

4.2 ผลการประเมิน

การทดสอบประสิทธิภาพของระบบ โดยผู้เชี่ยวชาญ (ช่างซ่อมรถ) จำนวน 5 คน และผู้ใช้ทั่วไป จำนวน 15 คน ด้วยแบบประเมินที่สร้างขึ้น การวิเคราะห์ข้อมูลทางด้านผู้เชี่ยวชาญ มีค่าเฉลี่ยที่ได้เท่ากับ 4.04 (ค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.18) และผู้ใช้ทั่วไป มีค่าเฉลี่ยที่ได้เท่ากับ 4.20 (ค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.17) ดังนั้นสามารถสรุปผลการประเมินความพึงพอใจของระบบโดยรวมได้ว่าระบบสนับสนุนการตัดสินใจสำหรับพิจารณาตรวจสอบอาการและบำรุงรักษารถยนต์เบื้องต้น(Basic Automotive Fault Diagnosis and Repair System) แก่มีความพึงพอใจอยู่ในระดับดีมาก เนื่องจากระบบสามารถประมวลผลและแสดงผลลัพธ์ได้ตรงตามเงื่อนไขและตรงตามความต้องการของผู้ใช้ ดังตารางที่ 2

ตารางที่ 2 : สรุปการประเมินประสิทธิภาพแอปพลิเคชัน

ด้าน	ผู้เชี่ยวชาญ (ช่างซ่อมรถ)		ผู้ใช้ทั่วไป	
	\bar{X}	S.D.	\bar{X}	S.D.
1.ความสามารถทำงานตามความต้องการของผู้ใช้	4.20	0.22	4.27	0.16
2.ด้านหน้าของแอปพลิเคชัน	3.80	0.22	4.07	0.12
3.ด้านการใช้งานของแอปพลิเคชัน	4.40	0.27	4.27	0.21
4.ด้านประสิทธิภาพของแอปพลิเคชัน	3.80	0.22	4.07	0.19
5.ด้านความปลอดภัยของแอปพลิเคชัน	4.00	0.00	4.33	0.19
ค่าเฉลี่ยโดยรวมทุกด้าน	4.04	0.19	4.20	0.17

5. ผลการทดลอง และข้อเสนอแนะ

5.1 สรุปผลการทำงานของระบบ

จากการพัฒนาแอปพลิเคชันได้มีการทดลอง และทดสอบประสิทธิภาพของโปรแกรมโดยใช้แบบประเมินความพึงพอใจจากกลุ่มตัวอย่าง 2 กลุ่ม ได้แก่ กลุ่มผู้เชี่ยวชาญ (ช่างซ่อมรถ) จำนวน 5 คน และกลุ่มผู้ใช้ทั่วไปจำนวน 15 คน โดยพบว่าผู้เชี่ยวชาญ (ช่างซ่อมรถ) มีความพอใจในแอปพลิเคชันอยู่ในระดับ ดีมาก เนื่องจากระบบสามารถประมวลผลและแสดงผลลัพธ์ได้ตรงตามเงื่อนไขและตรงตามความต้องการของผู้ใช้ โดยได้ค่าเฉลี่ยเท่ากับ 4.04 และส่วนเบี่ยงเบนมาตรฐานอยู่ในระดับ 0.18 ในส่วนของผู้ใช้ทั่วไปมีความพอใจในแอปพลิเคชันอยู่ในระดับ ดีมากเนื่องจากระบบสามารถประมวลผลและแสดงผลลัพธ์ได้ตรงตามเงื่อนไขและตรงตามความต้องการของผู้ใช้ โดยได้ค่าเฉลี่ยเท่ากับ 4.20 และส่วนเบี่ยงเบนมาตรฐานอยู่ในระดับ 0.17

สรุปได้ว่า แอปพลิเคชันตรวจสอบอาการและบำรุงรักษารถยนต์เบื้องต้นมีความพอใจอยู่ในระดับที่ดีมาก และสามารถนำไปใช้งานได้จริง

5.2 ข้อเสนอแนะ

เพื่อให้แอปพลิเคชันตรวจสอบอาการและบำรุงรักษารถยนต์เบื้องต้นสามารถมีประสิทธิภาพมากยิ่งขึ้นจึงเห็นควรที่จะได้รับการพัฒนาดังต่อไปนี้

5.2.1 ในส่วนของผู้เชี่ยวชาญเห็นว่า ควรจะเพิ่มข้อมูลที่ครอบคลุมมากยิ่งขึ้น เพื่อเพิ่มศักยภาพในการใช้แอปพลิเคชันมากยิ่งขึ้น อันจะก่อให้เกิดประโยชน์และประสิทธิภาพสูงสุด

5.2.2 โมเดลของการสนับสนุนการตัดสินใจ อาจมีการเพิ่มเติม ซึ่งกฎเกณฑ์ต่างๆเหล่านี้เป็นกฎที่ได้มาจากผู้เชี่ยวชาญ แต่ยังสามารถปรับปรุงและเพิ่มเติมได้ เนื่องจากอาจมีการได้รับข้อมูลที่เข้ามามากขึ้น จึงควรมีการปรับโมเดลของการสนับสนุนการตัดสินใจ ให้เหมาะสมตามความต้องการของผู้ใช้

เอกสารอ้างอิง

- [1] John Durkin, *Rule-Based Expert Systems*. Macmillan Publishing Company :United States, 1994.
- [2] ดร.วิลาศ ววงศ์, *ระบบผู้เชี่ยวชาญ*, กรุงเทพมหานคร : สถาบันเทคโนโลยีพระจอมเกล้าธนบุรี, 2535.
- [3] การแสดงความรู้ด้วยกฎ .[Online]. 1998. Availablefrom: <http://teacher.en.rmutt.ac.th/ktw/Resources/Full%20paper%20PDF/Chapter5> [2013 December 10].
- [4] “What is a rule-based system?,” [Online]. 1996. Availablefrom: <http://www.jpaine.org/students/lectures/lect3/node5.html> [2013 December 10].
- [5] มหาวิทยาลัยสุโขทัยธรรมราช, “ระบบสนับสนุนการตัดสินใจ,” *ระบบผู้เชี่ยวชาญ*, นนทบุรี : มหาวิทยาลัยสุโขทัยธรรมราช, 2550.
- [6] วัฒนาวิจิตร กัดดี, *ไดอะแกรมการตัดสินใจ*. พิมพ์ครั้งที่ 1 . เชียงใหม่ : สำนักพิมพ์มหาวิทยาลัยเชียงใหม่, 2531.
- [7] ชาญภูมิ โอภาภิบาล, “แอนดรอยด์คืออะไร,” *คู่มือแอนดรอยด์ฉบับสมบูรณ์*, นนทบุรี : ไอดีซีฯ, 2554.
- [8] John Watkins, Simon Mills. *Testing IT : an off-the-shelf software testingprocess*. Cambridge : Cambridge University Press, 2010.
- [9] Anderson. *Likert Scales, Education Research Methodology and Measurement : An International Handbook*. Victoria : Pergamon, 1988.
- [10] Likert, Rensis. A Technique for the Measurement of Attitudes : *Archives of Psychology*. Americas : n.p., 1932.

การพยากรณ์ราคาประเมินที่ดินโดยใช้เทคนิคเหมืองข้อมูล Land Price Forecasting using Data Mining Techniques

กรรณิการ์ นุชชมพู (Kannika Nutchomphu)¹ และ มาลีรัตน์ โสदानิล (Maleerat Sodanil)²

^{1, 2} ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

kannika.trd@gmail.com¹, msn@kmutnb.ac.th²

บทคัดย่อ

งานวิจัยนี้เป็นการประยุกต์ใช้เทคโนโลยีเหมืองข้อมูล มาใช้ในการพยากรณ์ราคาประเมินที่ดิน เพื่อนำไปใช้เป็นแนวทางสำหรับการประเมินราคาที่ดิน และช่วยประกอบการตัดสินใจของผู้บริหารหรือการตรวจสอบราคาประเมินที่ดินในปัจจุบัน โดยทำการรวบรวมข้อมูลปัจจัย ตัวแปร เช่น ระยะความลึกมาตรฐาน ลักษณะการใช้ประโยชน์ที่ดิน ลักษณะผิวจราจร ข้อมูลราคาตลาด และข้อมูลราคาประเมิน ในพื้นที่อำเภอลาดหลุมแก้ว จังหวัดปทุมธานี โดยใช้เทคนิคการจัดหมวดหมู่ (Classification) มาทำการสร้างโมเดลเพื่อให้ได้ราคาประเมินที่ดินที่ใกล้เคียงกับราคาตลาดมากที่สุด ซึ่งจากผลการทดสอบปรากฏว่า ปัจจัยสำคัญที่มีผลทำให้ราคาประเมินแตกต่างกันส่วนมากเป็นเรื่องเกี่ยวกับผิวจราจร ไม่ว่าจะเป็นความกว้างของผิวจราจร หรือประเภทของผิวจราจร ซึ่งมีระดับความสำคัญรวมกันถึง 39% และพบว่าทฤษฎีที่ให้ราคาประเมินที่ดีที่สุด คือ Generalized Linear และ Neural Network ซึ่งให้ค่าความแม่นยำถึง 97% ส่วน Decision Tree ก็ให้ค่าความแม่นยำที่สามารถยอมรับได้คือ 86% ซึ่งสามารถนำไปใช้ประโยชน์สำหรับการสร้างระบบเพื่อการประเมินราคาที่ดินคราวละมากเปลี่ยนแปลง (Mass Appraisal) หรือ ระบบแบบจำลองช่วยประเมินราคาคราวละมากเปลี่ยนแปลง (CAMA: Computer Assisted Mass Appraisal)

คำสำคัญ: เหมืองข้อมูล การจัดหมวดหมู่ข้อมูล การพยากรณ์ข้อมูล การประเมินราคาที่ดิน การประเมินราคาที่ดินคราวละมากเปลี่ยนแปลง

Abstract

This research is an applied of data mining technology. Used to forecasting value of land. To be used as a guide for land valuation. And the decision of the management or monitoring appraised land factors record. Data collected by variable factors such as the depth of the plot. Characteristics, land use Appearance barrier Type of road surface Market price information And appraisal Lat Lum Kaeo district in Pathumthani Using techniques for classification. to create models for the valuation of land that is close to market value as possible. The results of the test appear. A key factor that resulted in many different appraisal is about road surface. Whether it is the width of the road surface. Or the type of road surface Which have priority as to 39% and found that the theory that the appraisal is best to Generalized Linear and Neural Network, which provides accuracy up to 97% of the Decision Tree is an accuracy acceptable. was 86%, which can be utilized for building the system to convert more land valuation period (Mass Appraisal) system model helps estimate the conversion period (CAMA: Computer Assisted Mass Appraisal).

Keyword: Data Mining, Classification, Prediction, Land Price, Mass Appraisal

1. บทนำ

การประเมินราคาทรัพย์สิน มีความสำคัญและเกี่ยวข้องกับการทำธุรกรรมอย่างกว้างขวาง ไม่ว่าจะเป็นการทำธุรกรรมในตลาดเงิน ตลาดทุนหรือแม้กระทั่งระบบเศรษฐกิจของประเทศ เช่น การประเมินมูลค่าทรัพย์สินของบริษัทด้านการพัฒนาอสังหาริมทรัพย์เพื่อทำการเสนอขายหลักทรัพย์ต่อประชาชน การประเมินราคาทรัพย์สินเพื่อการเช่า ซื้อ ขาย หรือเพื่อการจำนองไว้เป็นหลักประกัน การประเมินราคาสำหรับการโอนหลักทรัพย์เพื่อชำระหนี้ หรือแม้กระทั่งการประเมินราคาเพื่อการร่วมทุน การเวนคืน การแบ่งแยกมรดกเพื่อความ เป็นธรรมแก่ลูกหลานและเป็นฐานในการคำนวณค่าธรรมเนียม ในการจดทะเบียนสิทธิและนิติกรรมเมื่อมีการซื้อ ขาย จำนอง ขายฝาก โอนมรดก ฯลฯ ของภาครัฐ เป็นต้น

รูปแบบการประเมินราคาที่ดินในปัจจุบันนั้นถึงแม้ว่าจะมีการกำหนดมาตรฐานการประเมินราคาที่ดินเพื่อใช้เป็นแนวทางในการดำเนินงานไม่ว่าจะเป็นจากหน่วยงานภาครัฐ รัฐวิสาหกิจ หรือภาคเอกชน เช่น มาตรฐานการประเมินราคาทรัพย์สิน กรมธนารักษ์ มาตรฐานและจรรยาบรรณวิชาชีพการประเมินมูลค่าทรัพย์สินในประเทศไทย สำนักงานคณะกรรมการกำกับหลักทรัพย์และตลาดหลักทรัพย์ มาตรฐานการประเมินราคาทรัพย์สิน มูลนิธิประเมินค่าทรัพย์สินแห่งประเทศไทย เป็นต้น แต่ยังมีแปลงที่ดินอีกจำนวนมากที่ต้องอาศัยความรู้ความสามารถ ประสบการณ์หรือดุลพินิจในการกำหนดราคาประเมินให้ถูกต้อง แม่นยำและสอดคล้องกับราคาตลาด อีกทั้งจำนวนแปลงที่ดินที่ต้องทำการประเมินราคาในปัจจุบันมีจำนวนมากแต่จำนวนบุคลากรด้านประเมินราคาไม่เพียงพอ และวิธีการประเมินราคาที่ดินที่เปลี่ยนแปลงตามปกติอาจทำให้ต้องสูญเสียเวลาและงบประมาณจำนวนมากในการดำเนินงาน นอกจากนี้อาจก่อให้เกิดความไม่เป็นธรรมต่อเจ้าของที่ดินหรือไม่สามารถอธิบายราคาประเมินที่เปลี่ยนแปลงไปตามลักษณะปัจจัยของที่ดินได้เนื่องจากมีการใช้ดุลพินิจในการประเมินราคา

การนำเทคโนโลยีเหมือนข้อมูลมาประยุกต์ใช้ในการประเมินราคาที่ดินนั้น จะช่วยให้สามารถประเมินราคาที่ดินได้คราวละมากแบบ (Mass Appraisal) และได้ราคาประเมินที่ดินที่เหมาะสมและใกล้เคียงกับราคาตลาดมากที่สุด ถึงแม้ว่าแปลง

ที่ดินจะมีลักษณะทางกายภาพ สภาพแวดล้อมที่แตกต่างกัน ซึ่งเป็นลักษณะการทำงานที่มูลค่าหรือราคาประเมินขึ้นอยู่กับปัจจัย ตัวแปรต่างๆ ที่เกี่ยวข้อง เช่น ราคาประเมินที่ดินขึ้นอยู่กับสภาพการใช้ประโยชน์ที่ดิน ระยะห่างจากถนน ฯลฯ การนำเทคนิคเหมือนข้อมูลมาใช้ในการสร้างแบบจำลองจะทำให้สามารถคำนวณได้ว่าถ้าแปลงที่ดินที่มีระยะห่างจากถนนลดหลั่นกันไป แปลงละ 20 เมตรจะมีราคาประเมินที่ดินเป็นเท่าใด ซึ่งจะช่วยให้ประหยัดเวลาและงบประมาณในการประเมินราคาเป็นอย่างมาก

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 หลักการประเมินราคาทรัพย์สิน

การประเมินราคาทรัพย์สิน [1] หมายถึง การกำหนด การประมาณราคาทรัพย์สิน โดยผู้ที่มีความรู้ มีประสบการณ์ว่าทรัพย์สินควรมีราคาหรือมูลค่าเท่าไรในช่วงเวลาใดเวลาหนึ่ง ซึ่งส่วนใหญ่เป็นการประมาณจากราคาตลาดที่มีการซื้อขายกันอย่างเป็นธรรม โดย [2] ได้มีการกำหนดหลักเกณฑ์และวิธีการประเมินราคาตามรูปแบบของทรัพย์สินและวัตถุประสงค์ของการประเมินราคา ได้แก่

1. วิธีเปรียบเทียบราคาตลาด (Market Approach)
2. วิธีพิจารณาจากรายได้ (Income Approach)
3. วิธีคิดจากต้นทุน (Cost Approach)
4. วิธีคำนวณมูลค่าคงเหลือสุทธิ (Residual Approach)

ซึ่งปัจจุบันทั้งภาครัฐและเอกชนส่วนใหญ่ใช้วิธีเปรียบเทียบราคาตลาด โดยราคาตลาดที่นำมาใช้นั้น [3] กำหนดให้สำรวจข้อมูลในท้องตลาดย้อนหลังไม่เกิน 3 ปีและจำเป็นต้องมีการตรวจสอบข้อมูลเบื้องต้น โดยการคำนวณหาค่า A/S Ratio (Assessed Value/Sales Ratio) เพื่อทราบทิศทางการเปลี่ยนแปลงราคาประเมินเบื้องต้น คือ

A/S Ratio = 1 หมายความว่าราคาประเมินเท่ากับราคาตลาด

A/S Ratio > 1 หมายความว่าราคาประเมินสูงกว่าราคาตลาด

A/S Ratio < 1 หมายความว่าราคาประเมินน้อยกว่าราคาตลาด

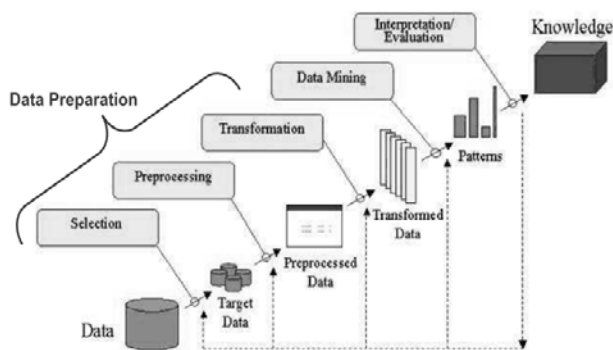
ซึ่งค่า A/S Ratio นั้นสามารถสะท้อนสถานะเศรษฐกิจในปัจจุบันว่าอยู่ในช่วงเศรษฐกิจขาขึ้น เศรษฐกิจขาลง หรืออยู่ในสภาวะทรงตัว ดังแสดงในตารางที่ 1

ตารางที่ 1: แสดงความสัมพันธ์ระหว่าง A/S Ratio กับสถานะเศรษฐกิจ

A/S Ratio	แนวโน้มการเพิ่มขึ้นของราคาที่ดิน	สถานะเศรษฐกิจ
0.3-0.8	25-30 %	เศรษฐกิจขาขึ้น
0.3-0.9	20-25 %	เศรษฐกิจทรงตัว
0.3-1.0	15-20 %	
0.3-1.1	10-15 %	
0.3-1.2	5-10 %	เศรษฐกิจขาลง
0.3-1.3	1-5 %	

2.2 การทำเหมืองข้อมูล (Data Mining)

การทำเหมืองข้อมูล [4] คือ กระบวนการค้นหาสารสนเทศหรือความรู้ที่อยู่ในฐานข้อมูลขนาดใหญ่ เพื่อนำความรู้ที่ได้ไปใช้ประโยชน์ในการตัดสินใจ โดยสารสนเทศที่ได้อาจนำมาพยากรณ์หรือสร้างต้นแบบสำหรับการจำแนก โดย [5] ขั้นตอนของการทำเหมืองข้อมูล ประกอบไปด้วย 1) การรวบรวมและคัดเลือกข้อมูลที่จะนำมาใช้ 2) การทำความสะอาดและเตรียมข้อมูลให้เหมาะสมแก่การนำไปใช้ 3) การปรับเปลี่ยนรูปแบบของข้อมูล เช่น การเชื่อมตารางในฐานข้อมูล เป็นการกำหนดโครงสร้างและลักษณะของข้อมูลที่ใช้ในการประมวลผล 4) การสร้างแบบจำลอง การเลือกเทคนิคที่เหมาะสมที่สุดเพื่อทำเหมืองข้อมูล ซึ่งประกอบไปด้วยเทคนิค ดังนี้ การหาความสัมพันธ์ของข้อมูล (Association Rule Discovery) การจัดกลุ่มและการทำนาย (Classification & Prediction) การแบ่งกลุ่ม (Clustering) และ การหาความแตกต่าง (Deviation Detection) จากนั้นจึงทำการทดสอบความถูกต้องและประสิทธิภาพของแบบจำลอง และ 5) การประเมินผลการทดลอง ซึ่งสามารถแสดงขั้นตอนดังรูปภาพที่ 1

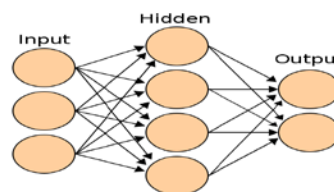


ภาพที่ 1: ขั้นตอนการทำงานของ Data Mining

ในส่วนของเครื่องมือและเทคโนโลยีที่ใช้ในการทำเหมืองข้อมูล มีดังนี้

2.2.1 เครือข่ายประสาท (Neural Network)

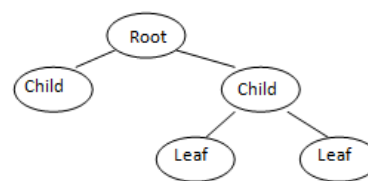
เป็นเทคโนโลยีที่มีพื้นฐานมาจากการจำลองการทำงานของเซลล์สมองของมนุษย์เพื่อใช้ในการคำนวณค่าฟังก์ชันจากกลุ่มข้อมูล โดยมีโครงสร้างเป็นกลุ่มของ Node ที่เชื่อมโยงถึงกันในแต่ละ Layer คือ Input Layer, Hidden Layer และ Output Layer ซึ่งสามารถแสดงได้ดังภาพที่ 2



ภาพที่ 2: โครงสร้างเครือข่ายประสาท

2.2.2 ต้นไม้ตัดสินใจ (Decision Tree)

เป็นการนำข้อมูลมาสร้างแบบจำลองการพยากรณ์ในรูปแบบโครงสร้างต้นไม้ตัดสินใจ ซึ่งเป็นลักษณะการทำงานแบบการเรียนรู้แบบมีครูสอน (Supervised Learning) จากกลุ่มของข้อมูลที่กำหนดไว้ล่วงหน้า (Traning Test) และสามารถพยากรณ์กลุ่มของข้อมูลที่ไม่เคยนำมาจัดหมวดหมู่ได้ โดยรูปแบบของต้นไม้ตัดสินใจประกอบด้วย โหนดแรกสุดเรียกว่า โหนดราก (Root Node) โหนดลูก (Child Node) และโหนดใบ (Leaf Node) ดังภาพที่ 3



ภาพที่ 3: โครงสร้างต้นไม้ตัดสินใจ

2.2.3 Generalized Additive Models (GAM)

พัฒนามาจาก Linear Regression และ Logistic Regression มีการตั้งสมมุติฐานว่า Model สามารถเขียนออกมาได้ในรูปแบบของผลรวมของ Possibly Non-Linear Function ซึ่งสามารถนำมาใช้ได้กับปัญหาแบบ Regression และ Classification

2.3 แบบจำลองเพื่อการประเมินราคา (CAMA)

แบบจำลองเพื่อการประเมินราคา (CAMA: Computer - Assisted Mass Appraisal) หรือ AVM (Automated Valuation Method) เป็นลักษณะของการประเมินราคาโดยอาศัยแบบจำลองทางสถิติซึ่งเป็นเทคนิคที่นำมาใช้ในการประเมินราคาคราวละมากแปลง (Mass Appraisal) ทำให้ประหยัดเวลาและต้นทุนในการประเมินราคาและช่วยลดอคติของผู้ประเมินราคา โดยส่วนใหญ่นำไปประยุกต์ใช้ในการจัดเก็บภาษีทรัพย์สินตามกฎหมายของรัฐหรือท้องถิ่นหรือการเวนคืน

ตามมาตรฐานการประเมินราคาทรัพย์สิน [6] กำหนดขั้นตอนของการประเมินโดยแบบจำลอง ดังนี้

1. การกำหนดคอสต์หรือทรัพย์สินที่พึงประเมิน เช่น เป็นที่อยู่อาศัย ที่เกษตรกรรมหรืออสังหาริมทรัพย์อื่นใด
2. การกำหนดขอบเขตพื้นที่ เช่น ในเขตเมือง
3. การกำหนดเงื่อนไข ปัจจัยหรือตัวแปรที่มีผลต่อมูลค่าตลาดของทรัพย์สิน
4. การพัฒนาโครงสร้างของแบบจำลองทางสถิติที่สะท้อนความสัมพันธ์เชิงตัวแปรต่อมูลค่าตลาดของทรัพย์สิน
5. การตรวจวัดความแม่นยำของตัวแปรที่มีผลต่อการเปลี่ยนแปลงมูลค่าทรัพย์สิน
6. การนำแบบจำลองไปปรับใช้ในพื้นที่ เพื่อขยายผลให้สามารถประเมินค่าทรัพย์สินได้อย่างประหยัด ถูกต้อง และรวดเร็ว
7. การปรับปรุงแบบจำลองอย่างต่อเนื่องเพื่อความเหมาะสมกับภาวะตลาดยิ่งขึ้นในอนาคต

ดังนั้นการทำ CAMA จะได้ผลดีหรือไม่จึงขึ้นอยู่กับข้อกำหนด และดำเนินการตามขั้นตอนข้างต้น โดย [7] คุณลักษณะที่สำคัญในการประเมินราคาที่ดิน จำเป็นต้องทำการแยกมูลค่าของอาคารสิ่งปลูกสร้างออกจากมูลค่าทรัพย์สินเพื่อให้ได้ราคาประเมินที่ดินที่แท้จริง ซึ่งมูลค่าของที่ดินนั้นต้องทำการตรวจคุณภาพและกำหนดราคาใหม่อย่างน้อยทุก 4 ปี ถึง 6 ปี

2.4 งานวิจัยที่เกี่ยวข้อง

ในส่วนของเทคนิคที่มีการนำมาใช้ในการสร้างแบบจำลองเพื่อทำการพยากรณ์ข้อมูลต่างๆ และการประเมินราคานั้นมีงานวิจัยที่เกี่ยวข้อง ดังนี้

อัมรินทร์และสมจิตร [8] ได้ทำการออกแบบและพัฒนาโมเดลการพยากรณ์ราคาข้าวเปลือกล่วงหน้า จำนวน 4 ชนิด จากการวัดประสิทธิภาพของโมเดลพบว่าค่าความถูกต้องมากที่สุด คือ 99.91% น้อยสุด คือ 77.01% ซึ่งสามารถนำไปใช้ประโยชน์แก่เกษตรกรผู้ปลูกข้าวได้ และเรวดี [9] ได้นำ Data Mining มาทำการทำนายผลสัมฤทธิ์ทางการเรียนของนักศึกษาตามพฤติกรรมทางการเรียนของนักศึกษาในแต่ละสาขาวิชา โดยใช้เทคนิค Decision Tree ในการจัดกลุ่มข้อมูล (Classification) และใช้ Neural Network ในการพยากรณ์/ทำนายข้อมูล (Predictive)

ในส่วนของการนำ Data Mining มาช่วยในการประเมินราคานั้น Claudioi, Vincenzo and Ruggiero [10] ได้นำเทคนิคต้นไม้ตัดสินใจ (Decision Tree) และรูปแบบการถดถอย (Regression) มาสร้างแบบจำลองการพยากรณ์ราคาซื้อขายเพื่อหารูปแบบและความปฏิสัมพันธ์ระหว่างการพยากรณ์ราคาประเมินอสังหาริมทรัพย์ ส่วน Carlos [11] ได้ทำการวิจัยเกี่ยวกับการประเมินราคาบ้านพักอาศัย โดยใช้ Decision Tree (M5) Neural Network และ Linear Regression (MLP) ซึ่งจากการวิจัยสรุปว่าได้ Decision Tree เหมาะสำหรับพื้นที่เขตเมือง ส่วน Neural Network (ANN) และ Linear Regression (MLP) เหมาะสำหรับพื้นที่ที่เป็นลักษณะกว้างๆ ทั่วไป แต่ Steven [12] ได้ประเมินราคาที่ดินโดยใช้ Neural Network (ANN) แต่ผลที่ได้มีความผิดพลาดมากเนื่องจากจำนวนและประเภทของตัวแปรที่ใช้และการจัดกลุ่มตัวอย่างไม่เหมาะสม อาทิ หน่วยที่วัดไม่เหมือนกัน เช่น เป็นฟุต เป็นเมตร เป็นเอเคอร์ ส่วนในประเทศใต้หวัน Lai Pi [13] ได้นำ Neural Network มาใช้ในการทำ CAMA สำหรับประเมินราคาทรัพย์สินเพื่อใช้ในการจัดเก็บภาษีและการจองเงินกู้ ซึ่งได้ประสิทธิภาพที่ดีกว่าการใช้แบบจำลองการถดถอยแบบดั้งเดิม

3. วิธีดำเนินการวิจัย

วิธีการวิจัยแบ่งออกเป็น 3 ขั้นตอน คือ การเตรียมข้อมูล การสร้างและทดสอบแบบจำลองการพยากรณ์ และการประเมินความถูกต้องและวัดประสิทธิภาพ

3.1 การเตรียมข้อมูล

ข้อมูลที่ใช้ในงานวิจัยประกอบไปด้วยการรวบรวมข้อมูลแปลงที่ดินจากฐานข้อมูลทางทะเบียนที่ดิน ข้อมูลการสำรวจปัจจัย /ตัวแปรจากการเดินสำรวจภาคสนาม ข้อมูลราคาตลาด

เบื้องต้นจากการจดทะเบียนสิทธิและนิติกรรม โดยใช้ข้อมูลในพื้นที่อำเภอลาดหลุมแก้ว จังหวัดปทุมธานี จำนวน 30,000 แปลง และทำการคัดเลือกปัจจัยที่สนใจนำมาสร้างแบบจำลองการพยากรณ์ ดังตารางที่ 2

ตารางที่ 2: ปัจจัยที่ใช้ในการพยากรณ์

ลำดับ	รายละเอียดปัจจัย
1	ตำแหน่งแปลงที่ดิน
2	การใช้ประโยชน์ของแปลงที่ดิน
3	ความลึกของแปลงที่ดิน
4	ประเภทผิวจราจร
5	ความกว้างผิวจราจร
6	ประเภทถนน
7	ความสะดวกในการเข้า-ออก
8	แปลงที่ดินติดถนน
9	แปลงที่ดินเป็นถนนหรือทาง
10	ราคาประเมินเดิม
11	ราคาตลาด

จากนั้นจึงทำการกรองข้อมูล (Data Cleaning) เพื่อทำความสะอาดข้อมูลให้มีคุณภาพและสมบูรณ์ เช่น การเติมข้อมูลที่ขาดหายหรือไม่สอดคล้องกัน และทำการแปลงรูปแบบข้อมูล (Data Transformation) ให้อยู่ในรูปแบบที่เหมาะสมสำหรับการนำไปสร้างและทดสอบด้วยโปรแกรม SPSS Modeler

3.2 การสร้างโมเดล (Modeling)

วิธีการสร้างตัวแบบการพยากรณ์ในงานวิจัยนี้เลือกใช้โปรแกรม IBM SPSS Modeler 15 ซึ่งเป็นซอฟต์แวร์ที่ได้พัฒนาขึ้นเพื่อสนับสนุนการสร้าง Model ทำงานในลักษณะของ User Interface มาช่วยในการสร้างแบบจำลองตามกฎการจัดหมวดหมู่ (Classification) โดยใช้เทคนิคเพื่อทำการเปรียบเทียบจำนวน 3 วิธีคือ (1.) ต้นไม้ตัดสินใจ (Decision Tree) ซึ่งใช้อัลกอริทึม CART (2.) เครือข่ายประสาทเทียม (Neural Network) และ (3.) Generalized Linear โดยทำการแบ่งข้อมูลออกเป็น 2 ส่วน คือ ข้อมูลชุดเรียนรู้ (Training set) และข้อมูลชุดทดสอบ (Testing set) โดยใช้วิธีการแบ่งแบบสุ่มแบบแยกค่าร้อยละ (Percentage Split) โดยกำหนดอัตราส่วนระหว่างชุดข้อมูลเรียนรู้กับชุดข้อมูลทดสอบเป็น 70% : 30% จากนั้นจึงทำ

การพัฒนาตัวแบบจำลองการพยากรณ์เพื่อทำนายค่าของราคาประเมินที่ดินที่ต้องการ โดยกำหนดค่าความเชื่อมั่นไว้ที่ 90%

3.3 การประเมินความถูกต้อง (Accuracy) และวัดประสิทธิภาพ (Evaluation)

3.3.1 ระดับความใกล้เคียงราคาตลาด (Accuracy) ตัวชี้วัดคือ Mean A/S Ratio กรณีแจกแจงปกติ หรือ Median A/S Ratio กรณีแจกแจงไม่ปกติ แล้วทดสอบว่าค่ากลางของ A/S Ratio เข้าใกล้ 1 หรือไม่

3.3.2 ความสม่ำเสมอระดับความใกล้เคียงราคาตลาด (Uniformity) ตัวชี้วัดคือ Coefficient of Variation : COV กรณีข้อมูล A/S Ratio แจกแจงปกติ หรือใช้ Coefficient of Dispersion : COD กรณีข้อมูล A/S Ratio แจกแจงไม่ปกติ

ค่ามาตรฐานที่ยอมรับได้ cov หรือ cod < 20%

3.3.3 ความเป็นธรรม/อคติ (Bias) เป็นการทดสอบระดับความใกล้เคียงราคาตลาดระหว่างทรัพย์สินที่มีมูลค่าต่ำกับทรัพย์สินที่มีมูลค่าสูงว่ามีความสม่ำเสมอหรือไม่ เป็นการทดสอบความสัมพันธ์ระหว่าง A/S Ratio กับมูลค่าทรัพย์สิน (ราคาที่ดินต่อตารางวา)

4. ผลการดำเนินงาน

จากการพัฒนาแบบจำลองการพยากรณ์ราคาประเมินที่ดินปรากฏว่ามีประสิทธิภาพที่ดีมากเนื่องจากสามารถพยากรณ์ราคาประเมินได้ใกล้เคียงกับราคาตลาดด้วยความแม่นยำถึง 97 % และ 86% ซึ่งเทคนิคที่ได้ผลดีที่สุดคือ Generalized Linear และ Neural Network โดยสามารถสรุปได้ดังตารางที่ 3

ตารางที่ 3: ค่าประสิทธิภาพจากการทดสอบเทคนิคต่างๆ

Techniques	Correlation	Accuracy	Uniformity
Generalized Linear	97.7 %	0.9	6.35
Neural Network	97.4 %	0.9	9.0
Decision Tree	86.2 %	0.8	5.35

จากตารางที่ 3 จะสังเกตได้ว่าความสม่ำเสมอระดับความใกล้เคียงราคาตลาด (Uniformity) ของ Decision Tree จะ

พยากรณ์ราคาประเมินที่ระดับใกล้เคียงกันมากที่สุด กล่าวคือ ช่วงของราคาประเมินที่ดินที่ได้ไม่แตกต่างกันมากนัก

5. สรุปผลการทดลอง

การนำเหมืองข้อมูลมาประยุกต์ใช้เพื่อช่วยในการพยากรณ์ราคาประเมินที่ดินโดยใช้กฎการจัดหมวดหมู่ โดยใช้เทคนิค 3 วิธี ได้แก่ ต้นไม้ตัดสินใจ (Decision Tree) เครือข่ายประสาทเทียม (Neural Network) และ Generalized Linear โดยมีการแบ่งข้อมูลแบบสุ่มด้วยการแบ่งร้อยละของชุดเรียนรู้และชุดทดสอบออกจากกัน เมื่อทำการวัดค่าประสิทธิภาพของแบบจำลอง ได้แก่ ค่าความแม่นยำ ค่าความใกล้เคียงราคาตลาด ความสม่ำเสมอระดับความใกล้เคียงราคาตลาด และค่าความเป็นธรรมชาติ โดยมีค่าความแม่นยำมากที่สุดถึง 97% คือ เทคนิค Generalized Linear และ Neural Network ซึ่งมีความถูกต้องแม่นยำสูงและเหมาะสมกว่าเทคนิค Decision Tree ซึ่งหากมีการนำแบบจำลองการพยากรณ์ดังกล่าว ไปสร้างเป็นแอปพลิเคชันสำหรับการพยากรณ์ราคาประเมินที่ดินจะช่วยให้ได้ราคาประเมินที่สามารถสะท้อนราคาได้ดี และช่วยให้ลดระยะเวลาและค่าใช้จ่ายในการประเมินราคาที่ดินเป็นอย่างมาก แต่การนำไปใช้ให้ได้มีประโยชน์สูงสุดและถูกต้องที่สุดนั้น จะต้องมีความระมัดระวังในการเก็บข้อมูลปัจจัยต่างๆ โดยเฉพาะข้อมูลราคาตลาดจะต้องอาศัยการสอบถามเพื่อให้ได้ราคาที่ตรงกับความเป็นจริงที่สุด และถ้าหากมีการนำข้อมูลแผนที่ทางภูมิศาสตร์ (GIS) มาช่วยในการเดินสำรวจจะทำให้สามารถตรวจสอบและบันทึกข้อมูลได้แม่นยำมากขึ้นและสามารถเห็นภาพรวมของราคาประเมินตามระยะความลึกที่แตกต่างกันได้ชัดเจนยิ่งขึ้น

เอกสารอ้างอิง

- [1] ไพโรจน์ ชิงศิลป์, "หลักการประเมินราคาทรัพย์สิน", พิมพ์ครั้งที่ 2 ปรับปรุงใหม่. : โรงพิมพ์สุธาสิน, กรุงเทพฯ, 2538.
- [2] สำนักงานคณะกรรมการกำกับหลักทรัพย์และตลาดหลักทรัพย์, "มาตรฐานและจรรยาบรรณวิชาชีพการประเมินมูลค่าทรัพย์สินในประเทศไทย", 2548.
- [3] คณะกรรมการกำหนดราคาประเมินทุนทรัพย์. "ระเบียบคณะกรรมการกำหนดราคาประเมินอสังหาริมทรัพย์", หมวด 1 ข้อ 9, กระทรวงมหาดไทย, 2538.
- [4] สุชาติ ธีระนันท์, "การเพิ่มศักยภาพระบบการจัดการด้วยเทคโนโลยีสารสนเทศ", *สัมมนาหนึ่งทศวรรษไอทีอุทา, กรุงเทพฯ*, 2545.
- [5] ชิดชนก ส่งศิริ, ธนาวิรัตน์ รักธรรมานนท์ และกฤษณะ ไวยมัย, "การใช้เทคนิค Data Mining เพื่อค้นหาภาควิชาที่เหมาะสมที่สุดให้กับนิสิต", *การประชุมวิชาการของมหาวิทยาลัยเกษตรศาสตร์*, ครั้งที่ 39.
- [6] The Appraisal Foundation, "Uniform Standards of Professional Appraisal Practice", American United Nation, 2002
- [7] International Assessing Association Officers, "Standard on Mass Appraisal of Real Property" 2013.
- [8] อัมรินทร์ ก้อนแพง , สมจิตร อาจอินทร์, "การพยากรณ์ราคาข้าวเปลือกโดยใช้เทคนิคการทำเหมืองข้อมูล", *การประชุมวิชาการเสนอผลงานวิจัยระดับบัณฑิตศึกษาแห่งชาติ*, ครั้งที่ 23, 2550.
- [9] เรวดี สักดิ์คุณธรรม, "การใช้เทคนิคดาต้าไมนิ่งในการสร้างฐานความรู้ เพื่อการทำนายสัมฤทธิ์ผลทางการเรียนของนักศึกษาวิทยาลัยราชพฤกษ์", *กรุงเทพฯ*, 2542.
- [10] Claudio Acciani, Vincenzo Fucilli and Ruggiero Sardaro. "Data mining in real estate appraisal :a model tree and multivariate adaptive regression spline approach", 2011.
- [11] Carlos del Cacho, "A comparison of data mining methods for mass real estate appraisal", spane, 2010
- [12] Steven Peterson, "Neural Network Hedonic Pricing Models in Mass Real Estate Appraisal", *Virginia Commonwealth*
- [13] Lai Pi-ying, "Analysis of the Mass Appraisal Model by Using Artificial Neural Network in Kaohsiung City", *Journal of Modern Accounting and Auditing*, October 2011, Vol.7, No.10

การกระจายส่วนแบ่งของอาเซียนเพาเวอร์กริด ในโครงสร้างการรวมกลุ่มที่ดีที่สุดด้วยหลักของค่าแชปเปลีย์

Payoff Distribution of ASEAN Power Grid in Optimal Coalition Structure by Shapley Value

เบญจวรรณ อินทระ (Benjawan Intara)¹ พาณิชย์ สุกโคต (Panich Sodkhot)²

และฉัตรตระกูล สมบัติธีระ (Chattrakul Sombattheera)³

สาขาเทคโนโลยีสารสนเทศ คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม

¹Panglum_pp@hotmail.com, ²pon.ict@gmail.com, ³chattrakul.s@msu.ac.th

บทคัดย่อ

มูลค่าของแหล่งทรัพยากรในอาเซียนเพาเวอร์กริด (ASEAN Power Grid) สามารถนำมาเป็นใช้เป็นข้อมูลในการคำนวณหาส่วนแบ่งให้กับแต่ละประเทศที่เป็นสมาชิกของอาเซียนเพาเวอร์กริด โดยการใช้หลักคิดของอัลกอริทึมการรวมกลุ่มที่ดีที่สุดและอัลกอริทึมค่าแชปเปลีย์ในเรื่องการจัดสรรส่วนแบ่ง (Payoff) อย่างยุติธรรมของมูลค่าของแหล่งทรัพยากรในอาเซียน ในงานวิจัยนี้เราต้องการศึกษาส่วนแบ่งของอาเซียนเพาเวอร์กริดซึ่งผลลัพธ์ในการทดลองนี้ทำให้ทราบถึงลักษณะการกระจายส่วนแบ่งของอาเซียนเพาเวอร์กริดในแต่ละกลุ่มย่อยของโครงสร้างการรวมกลุ่มที่ดีที่สุด

คำสำคัญ: อาเซียนเพาเวอร์กริด ส่วนแบ่ง ค่าแชปเปลีย์ โครงสร้างการรวมกลุ่มที่ดีที่สุด

Abstract

Value of resource in ASEAN Power Grid. It is data for calculate payoff distribution among ASEAN Power Grid by Optimal Coalition Structure and Shapley value presents the concept of fairness to among ASEAN Power Grid. This research seeks to study the pattern of payoff distribution among ASEAN Power Grid.

Keywords: ASEAN Power Grid, payoff, Shapley value, Optimal Coalition Structure.

1. บทนำ

อาเซียนเพาเวอร์กริดเป็นแนวคิดในการเชื่อมโยงระบบไฟฟ้าและรวมกันจัดสรรการใช้ทรัพยากรพลังงานไฟฟ้าใน

ประเทศอาเซียน ที่เป็นการรวมพลังของทุกชาติในภูมิภาคส่งเสริมแนวคิดการใช้พลังงานอย่างยั่งยืนและมีประสิทธิภาพเพื่อความมั่นคงทางด้านพลังงานและเศรษฐกิจ รวมทั้งส่งเสริมให้อาเซียนพัฒนาทัดเทียมภูมิภาคอื่นของโลก และในปี 2012 มีนักเศรษฐศาสตร์คือ คุณลอยด์ แชปเปลีย์ ได้รับรางวัลโนเบลสาขาเศรษฐศาสตร์ ในการเสนอค่าแชปเปลีย์ที่เป็นทฤษฎีการจัดสรรอย่างยุติธรรม และออกแบบตลาดแนวใหม่ ซึ่งทำให้การจับคู่ความต้องการในตลาดมีประสิทธิภาพ

อย่างไรก็ตามในการศึกษาข้อมูลการวางแผนในเรื่องอาเซียนเพาเวอร์กริด ยังเป็นเรื่องการวางโครงข่ายการเชื่อมโยงระบบไฟฟ้า แต่ยังไม่มีการกล่าวถึงในเรื่องของการจัดสรรการกระจายส่วนแบ่งของอาเซียนเพาเวอร์กริด

ดังนั้นในงานวิจัยนี้จึงต้องการนำข้อมูลจริงของแหล่งพลังงานไฟฟ้าของอาเซียนเพาเวอร์กริดมาศึกษาลักษณะการกระจายตัวของขนาดของกลุ่มอาเซียนเพาเวอร์กริดกับมูลค่าของแหล่งทรัพยากรของอาเซียนที่ใช้ในการผลิตพลังงานไฟฟ้า โดยใช้หลักการของค่าแชปเปลีย์[1] ที่ได้รับการยอมรับในเรื่องของความยุติธรรม มาหาคำตอบในส่วนแบ่งสุดท้ายที่แต่ละประเทศจะได้รับ

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 เอเจนต์

เอเจนต์ (Agent)[2] คือ ระบบคอมพิวเตอร์ที่มีความสามารถในการกระทำบางสิ่งบางอย่างได้อย่างอิสระได้ด้วยตัวเอง ในนามของเจ้าของหรือผู้ใช้ ซึ่งเอเจนต์สามารถคิดวางแผนในการกระทำของตัวเองได้ หรือเรียนรู้จากสภาพแวดล้อม เพื่อให้บรรลุเป้าหมายตามที่ได้รับมอบหมาย

2.2 ระบบมัลติเอเจนต์

ระบบมัลติเอเจนต์ (Multi-agent system) [3, 4] คือ ระบบฯ หนึ่งที่ประกอบด้วยเอเจนต์จำนวนหนึ่ง ที่มีการปฏิสัมพันธ์ต่อกัน และสามารถทำงานตามคำสั่งที่ได้รับมอบหมาย เพื่อให้บรรลุเป้าหมายหรือตอบใจของผูู้ใช้งาน

2.3 ทฤษฎีเกม

ทฤษฎีเกม(Game theory)[5] แบบการร่วมมือเป็นการรวมกลุ่มของเอเจนต์ในเกมการร่วมมือที่สามารถเกิดขึ้นได้ในทุกๆบริบทที่ประกอบด้วยหน่วยการตัดสินใจ (Decision Making Unit)หลายๆหน่วย ทั้งนี้การรวมกลุ่มของเอเจนต์ยอมทำให้เกิดผลประโยชน์ร่วมกัน (Coalition Value) เกิดขึ้นในทางทฤษฎีนั้นผลประโยชน์ที่เกิดขึ้นจากการร่วมมือกันจะเป็นตัวเลขที่บ่งชี้มูลค่าหรือความพึงพอใจของเอเจนต์ต่อการรวมกลุ่มนั้นๆ ซึ่งอาจหมายถึงผลกำไร การลดต้นทุน ทั้งนี้สิ่งที่ทำให้เอเจนต์ซึ่งมีความเห็นแก่ตัว (Self Interested) ตกลงร่วมมือกันนั้นขึ้นอยู่กับการแข่งขันผลประโยชน์ที่จะเกิดจากการร่วมมือกันของสมาชิกในกลุ่ม หากสมาชิกในกลุ่มทั้งหมดมีความพอใจต่อส่วนแบ่งที่ตนจะได้รับจากผลประโยชน์ที่จะเกิดจากการร่วมมือกัน การรวมกลุ่มก็จะเกิดขึ้น

สมมุติให้ระบบหลายเอเจนต์ระบบหนึ่งประกอบด้วย 3 เอเจนต์ คือ $N = \{a, b, c\}$ ระบบนี้สามารถมีการรวมกลุ่ม ($S \subseteq N$) ได้ทั้งหมด $2^3 - 1 = 7$ กลุ่ม ได้แก่ $\{A\}, \{B\}, \{C\}, \{A, B\}, \{A, C\}, \{B, C\}$ และ $\{A, B, C\}$ จะเห็นได้ว่าการรวมกลุ่มสามารถเป็นได้ตั้งแต่เอเจนต์แยกตัวอยู่เดี่ยวๆ (Singleton Coalition) การรวมกลุ่มกันของสองเอเจนต์หรือมากกว่าไปจนกระทั่งถึงการรวมกลุ่มของทุกๆ เอเจนต์ (Grand Coalition) ในระบบ

ในทฤษฎีการรวมกลุ่มจะกำหนดให้มีฟังก์ชันกำหนดมูลค่าของกลุ่ม(Characteristic Function) ซึ่งจะทำหน้าที่กำหนดผลประโยชน์หรือมูลค่า ($U(s)$) ที่จะเกิดจากการรวมกลุ่มให้แก่กลุ่มต่างๆ ยกตัวอย่างเช่น

$$U(A) = 2, U(B) = 4, U(C) = 3,$$

$$U(A, B) = 5, U(A, C) = 6, U(B, C) = 8,$$

$$U(A, B, C) = 10$$

หมายถึงมูลค่ากลุ่มที่ประกอบด้วย $\{A\}, \{B\}, \{C\}, \{A, B\}, \{A, C\}, \{B, C\}$ และ $\{A, B, C\}$ ตามลำดับ ภายหลังจากการต่อรองส่วนแบ่งระหว่างกันแล้วในท้ายที่สุดแล้วเอเจนต์จะ

ตัดสินใจรวมกลุ่มกันอย่างไรอย่างหนึ่ง ลักษณะการรวมกลุ่มนั้นๆ เรียกว่า โครงสร้างการรวมกลุ่ม (Coalition Structure: CS) จะเห็นได้ว่าในระบบนี้จะมี CS ทั้งหมดอยู่ 5 แบบ ได้แก่ $\{A\}, \{B\}, \{C\}, \{A, B\}, \{C\}, \{A, C\}, \{B\}, \{A\}, \{B, C\}$ และ $\{A, B, C\}$ ซึ่งสามารถหามูลค่าของโครงสร้างการรวมกลุ่ม (Coalition Structure Value) ได้ดังนี้

$$U(CS) = \sum_{S \in CS} U(S) \quad (1)$$

สิ่งสำคัญในภาพรวมของระบบคือโครงสร้างการรวมกลุ่มที่ดีที่สุด (Optimal Coalition Structure) คือ

$$U(CS)^* = \operatorname{argmax} \sum_{S \in CS} U(S) \quad (2)$$

จากสมการที่ 2 แสดงรูปแบบโครงสร้างการรวมกลุ่มที่มีผลรวมของมูลค่ากลุ่มสูงสุด เช่น จากโครงสร้างการรวมกลุ่มทั้ง 5 ลักษณะที่มีมูลค่าดังนี้ $U(\{A\}, \{B\}, \{C\}) = 2 + 4 + 3 = 9$, $U(\{A, B\}, \{C\}) = 5 + 3 = 8$, $U(\{A, C\}, \{B\}) = 6 + 4 = 10$, $U(\{A\}, \{B, C\}) = 2 + 8 = 10$, $U(\{A, B, C\}) = 10$ ดังนั้นโครงสร้างการรวมกลุ่มที่ดีที่สุด (CS^*) คือ $\{A, C\}, \{B\}, \{A\}, \{B, C\}$ และ $\{A, B, C\}$ และ $U(CS)^*$ คือ 10

แนวคิดในการแก้ไขปัญหา (Solution Concept) ของการรวมกลุ่มในทฤษฎีเกมการร่วมมือมีหลายอย่างเช่น เคอร์เนล (Kernel)[6] ใช้แนวคิดในเรื่องของความเสถียร, คอร์ (The Core)[7] ใช้แนวคิดในเรื่องของประสิทธิภาพ และค่าแชปเปลีย์ ใช้แนวคิดในเรื่องของความยุติธรรมเป็นทฤษฎีที่สามารถจัดสรรส่วนแบ่งให้กับเอเจนต์ตามลำดับความสำคัญและมูลค่าที่เอเจนต์นำมามอบให้กับกลุ่ม โดยมีหลักการดังนี้

2.4 โครงการด้านพลังงานภายใต้ประชาคมเศรษฐกิจ

อาเซียน

เป็นแนวคิดด้านการนำแหล่งทรัพยากรของแต่ละประเทศในภูมิภาคอาเซียนมาใช้ให้เกิดประโยชน์สูงสุด โดยการเชื่อมโยงระบบไฟฟ้าในอาเซียน จึงช่วยลดปัญหาการซื้อเชื้อเพลิงจากกลุ่มประเทศนอกอาเซียน เพราะอาเซียนมีทรัพยากรด้านพลังงานที่หลากหลาย และเป็นการรวมพลังของทุกชาติในภูมิภาคอาเซียนที่จะร่วมกันจัดสรรทรัพยากรที่สามารถใช้พลังงานให้เกิดประโยชน์คุ้มค่ามากที่สุด

2.5 The Power Trading Agent Competition

จากการศึกษางานวิจัยนี้ Ketter และคณะ[8] ได้เสนอการจำลองการทำงานของเพาเวอร์แท็ค (Power TAC) ถูกออกแบบ

มาเพื่อใช้เป็นแบบจำลองการแข่งขันที่มีรูปแบบเสรีในตลาดค้าปลีกและค้าส่งของพลังงานไฟฟ้าระหว่างโบร์กเกอร์ ที่จะต้องแข่งขันหาลูกค้าที่ต้องการใช้งานพลังงานไฟฟ้าและผู้ประกอบการที่ผลิตพลังงานไฟฟ้า ซึ่งมีการจำลองสภาพแวดล้อมรูปแบบของการตลาดขายส่งพลังงานไฟฟ้า การควบคุมการกระจายของผลประโยชน์และจำนวนประชาชนที่เป็นลูกค้า จะมีโบร์กเกอร์เป็นผู้ขายส่งและขายปลีกพลังงานไฟฟ้า รวมทั้งแบบจำลองของผู้ประกอบการที่มีความหลากหลายของหน่วยองค์กรเชิงพาณิชย์และอุตสาหกรรมหลายแห่งที่กำลังการผลิต

3. วิธีดำเนินการวิจัย

งานวิจัยนี้ต้องการศึกษาลักษณะการกระจายตัวของส่วนแบ่งตามขนาดของกลุ่ม โดยการนำข้อมูลจริงของอาเซียนเพาเวอร์กริด (ASEAN Power Grid) ที่มีทรัพยากรที่ใช้ในการผลิตไฟฟ้าในภูมิภาคอาเซียนมาคิดคำนวณในการจัดสรรส่วนแบ่งด้วยกระบวนการของอัลกอริทึมค่าแฮปปีลีย์

3.1 ข้อมูลจริงของอาเซียนเพาเวอร์กริด

ข้อมูลสมาชิกในกลุ่มประเทศที่เป็นสมาชิกของภูมิภาคอาเซียนที่มีแหล่งทรัพยากรธรรมชาติที่ใช้ในการผลิตพลังงานไฟฟ้า เช่น พลังงานน้ำมัน ก๊าซธรรมชาติ ถ่านหิน และพลังงานน้ำ แสดงดังตารางที่ 1

ตารางที่ 1: แหล่งทรัพยากรที่ใช้ในการผลิตไฟฟ้าของภูมิภาคอาเซียน [9]

สมาชิกในกลุ่มประเทศอาเซียนเพาเวอร์กริด	น้ำมัน (ล้านบาร์เรล)	ก๊าซธรรมชาติ (ล้านล้านลูกบาศก์ฟุต)	ถ่านหิน (ล้านตัน)	พลังน้ำ (เมกะวัตต์)
อินโดนีเซีย	3,750	112.5	5,529	245,091
มาเลเซีย	5,357	82.4	4	26,256
บรูไน	1,200	12.4	-	-
เวียดนาม	4,700	7.7	150	34,247
ไทย	453	12.0	1,239	2,055
พม่า	50	20.8	2	39,726
ฟิลิปปินส์	138	3.3	316	5,365
ลาว	-	-	503	26,598
กัมพูชา	-	-	-	10,046
สิงคโปร์	-	-	-	-

3.2 การออกแบบอัลกอริทึมที่เหมาะสมกับข้อมูลอาเซียนเพาเวอร์กริด

อัลกอริทึมนี้จะทำการหาโครงสร้างการรวมกลุ่มที่ดีที่สุด ตามกระบวนการโดยเริ่มจากการอธิบายตัวแปรต่างๆในอัลกอริทึมดังตารางที่ 2 และกระบวนการคำนวณหาการรวมกลุ่มภายใน OCS แสดงดัง ภาพที่ 1 ดังนี้

```

Algorithm 1: Optimal Coalition Structure(OCS)
1: set N = noa of ASEAN Power Grid
2: for each i from 1 to N do
3:   set s = 2N - 1
4:   for each s from S ⊆ N do
5:     set v
6:     v(CS) = ∑ s ∈ CS vs
7:   end for
8: end for
9: generate L(N)
10: let F[L] <- CS
11: for each v(CS) ∈ CS do
12:   v(CS)* = argmax ∑ s ∈ CS vs
13:   OCS = v(CS)*
14: end for
15: return OCS
    
```

ภาพที่ 1: อัลกอริทึม OCS

ตารางที่ 2: แสดงการอธิบายตัวแปรใน Algorithm 1

ตัวแปร	อธิบายตัวแปร
N	จำนวนของเอเจนต์
L	การแบ่ง Integer Partition(วิธีการแบ่งเลขจำนวนเต็ม)
S	การรวมกลุ่มทั้งหมด (S ⊆ N) = 2 ⁿ - 1
F	การตั้งค่า(Configuration) ในการรวมกลุ่ม
CS	โครงสร้างการรวมกลุ่ม
v(CS)	มูลค่าของโครงสร้างการรวมกลุ่ม (Coalition Structure Value)
v(CS)*	มูลค่าของโครงสร้างการรวมกลุ่มที่ดีที่สุด
OCS	โครงสร้างการรวมกลุ่มที่ดีที่สุด (Optimal Coalition Structure)

จาก Algorithm 1 มีกระบวนการทำงาน ดังนี้

1) กำหนดจำนวนของเอเจนต์ N จำนวนแสดงดังบรรทัดที่ 1 เช่น กำหนดให้ N = 7 จะได้โครงสร้างการรวมกลุ่ม N = {อินโดนีเซีย, มาเลเซีย, บรูไน, เวียดนาม, ไทย, พม่า, ฟิลิปปินส์}

2) กำหนดจำนวนของโครงสร้างการรวมกลุ่มทั้งหมด S = 2⁷ - 1 = 127 แสดงดังบรรทัดที่ 3

3) ทำการให้กำเนิดมูลค่าในแต่ละ S พร้อมทั้งกำหนดมูลค่าให้กับโครงสร้างการรวมกลุ่มทั้ง 15 กลุ่ม ดังบรรทัดที่ 5 แล้วหามูลค่าของโครงสร้างการรวมกลุ่ม ในบรรทัดที่ 6

4) ให้กำเนิดการแบ่งเลขจำนวนเต็ม (Integer Partition) ในบรรทัดที่ 9 เริ่มจากให้ N แทนเลขจำนวนเอเจนต์ การแบ่งของ N คือ การแทน N ด้วยผลรวมของเลขจำนวนเต็มกล่าวคือ $N = a_1 + \dots + a_N$ ผลรวมของ $a_1 + \dots + a_N$ เรียกว่า Parts ซึ่งเป็นส่วนหนึ่งของการแบ่งเลขจำนวนเต็ม กำหนดให้ $L(N)$ แทนจำนวนของการแบ่งของ N และ $L(N)$ เรียกว่า Partition Number หรือจำนวนที่เกิดจากการแบ่งเลขจำนวนเต็ม ยกตัวอย่างเช่น มีเอเจนต์อยู่ 4 เอเจนต์ในระบบจะแทนด้วย $L(4)$ เป็นจำนวนการแบ่งของ 4 เอเจนต์ คือ 4, 3+1, 2+2, 1+1+2, 1+1+1+1 ดังตัวอย่างตารางที่ 2

5) ในบรรทัดที่ 10 ให้ $F[L]$ เป็นการแบ่งโครงสร้างการรวมกลุ่มตาม Integer Partition โดยกำหนดให้แต่ละ Integer Partition แทน แต่ละการตั้งค่าซึ่งสามารถมีการรวมกลุ่มได้ทั้งหมด 15 กลุ่ม แสดงดังตารางที่ 3

6) คำนวณหาโครงสร้างการรวมกลุ่มที่ดีที่สุดของอาเซียนเพาเวอร์กริดในแต่ละ CS ที่มีผลรวมสูงสุดจะเป็น OCS of ASEAN Power Grid แล้วนำคำนวณใน Algorithm 2 ต่อไป

ตารางที่ 3: ตารางแสดงโครงสร้างการรวมกลุ่ม

Integer Partition	CS
L[7]	$F[L] = \{ \{A,B,C,D\} \}$
L[5+1+1]	$F[L] = \left\{ \begin{array}{l} \{B,C,D\}, \{A\} \\ \{A,C,D\}, \{B\} \\ \{A,B,D\}, \{C\} \\ \{A,B,C\}, \{D\} \end{array} \right.$
L[4+1+1+1]	$F[L] = \left\{ \begin{array}{l} \{A,B\}, \{C,D\} \\ \{A,C\}, \{B,D\} \\ \{A,D\}, \{B,C\} \end{array} \right.$
L[3+2+1+1]	$F[L] = \left\{ \begin{array}{l} \{A\}, \{B\}, \{C,D\} \\ \{A\}, \{C\}, \{B,D\} \\ \{A\}, \{D\}, \{B,C\} \\ \{B\}, \{C\}, \{A,D\} \\ \{B\}, \{D\}, \{A,C\} \\ \{C\}, \{D\}, \{A,B\} \end{array} \right.$
L[2+2+1+1+1]	$F[L] = \{ \{A\}, \{B\}, \{C\}, \{D\} \}$

3.3 อัลกอริทึมแชปลีย์ที่ใช้ในการทดสอบ

อัลกอริทึมที่นำมาใช้เพื่อคำนวณหาส่วนแบ่งของประเทศในภูมิภาคอาเซียนเพาเวอร์กริด (APG: ASEAN Power Grid) แสดงดังตารางที่ 4

ตารางที่ 4 : แสดงการอธิบายตัวแปรที่ใช้ในอัลกอริทึม

ตัวแปร	อธิบายตัวแปร
P	เซตของ permutation โดย $p \in P$ หรือกลุ่มเอเจนต์ในกลุ่ม p โดย $APG_j < APG_k$
APG_j	ประเทศที่เป็นสมาชิกใน ASEAN Power Grid ที่ถูกเรียงลำดับการเข้าร่วมกลุ่ม โดยที่ $1 \leq j$
APG_k	ประเทศที่เป็นสมาชิกใน ASEAN Power Grid ที่ถูกเรียงลำดับการเข้าร่วมกลุ่ม โดยที่ $k \leq n$
$S(APG_j)$	กลุ่มเก่าที่มีสมาชิกเรียงตาม permutation ถึง APG_j
$S(APG_k)$	กลุ่มใหม่ที่มีสมาชิกเรียงตาม permutation ถึง APG_k
$U(Sapg_j)$	มูลค่ากลุ่มเก่า
$U(Sapg_k)$	มูลค่ากลุ่มใหม่
$C_{apg_k}^P$	Contribution ของ apg_k ใน P
Avg	ค่าเฉลี่ยของกำไรที่ถูกจัดสรรอย่างยุติธรรม

Algorithm 2: Modifier Shapley Value

```

1: APG = argmax $\sum_{s \in apg}$ 
2: for each p  $\in$  P do
3:   Set  $APG_j = \text{null}$ 
4:   for each  $APG_j \in p$  do
5:     set  $S(APG_j) = S(APG_k) \cup APG_j$ 
6:      $C_{apg_k}^P = v(S(apg_k)) - v(S(apg_j))$ 
7:      $S(apg_j) = S(apg_k)$ 
8:   end for
9:   for  $APG_j \in APG$  do
10:     $C(apg_j, p) = \sum_{i=0}^p C_{apg_k}^P$ 
11:   end for
12:   Arg = 1/APG  $C(apg_j, p)$ 
13: end for
14: return Arg
    
```

ภาพที่ 2: อัลกอริทึม Modifier Shapley Value

Algorithm 2 แสดงขั้นตอนการทำงานในการคำนวณหาส่วนแบ่งของกลุ่มสมาชิกในภูมิภาคอาเซียนเพาเวอร์กริดที่มีแหล่งทรัพยากรที่ใช้ในการผลิตพลังงานไฟฟ้า โดยการใช้หลักการของแชปลีย์ ดังต่อไปนี้

1) การหาการเรียงลำดับของสมาชิกในกลุ่มประเทศอาเซียน เพาเวอร์กริด ซึ่งแสดงใน for บรรทัดที่ 2 อย่างเช่น กลุ่มประเทศอาเซียนเพาเวอร์กริดที่มีแหล่งทรัพยากรน้ำมัน ยกตัวอย่างที่พิจารณา 3 ประเทศ ประกอบด้วยประเทศ {อินโดนีเซีย, มาเลเซีย, บรูไน} จะได้การเรียงลำดับของกลุ่ม คือ $3! = 6$ รายการ หรือพิจารณาทั้ง 7 ประเทศ คือ {อินโดนีเซีย, มาเลเซีย, บรูไน, เวียดนาม, ไทย, พม่า, ฟิลิปปินส์} จะได้การเรียงลำดับของกลุ่ม คือ $7! = 5,040$ รายการ จากนั้นจึงกำหนดให้ประเทศในกลุ่ม APG ที่ได้เรียงเข้ากลุ่มแล้วมีค่าเป็น null

2) การคำนวณหาส่วนแบ่งให้กับสมาชิกตามหลักค่าเฉลี่ย โดย for ในบรรทัดที่ 4 สำหรับแต่ละการเรียงลำดับสมาชิกของกลุ่ม APG ภายใน for ของการเรียงลำดับจะวนลูปเพื่อทำการคำนวณหามูลค่าที่แต่ละประเทศในกลุ่ม APG_j นำมามอบให้กับกลุ่ม จากนั้นคำนวณหามูลค่าที่แต่ละประเทศในกลุ่ม APG_j นำมามอบให้กับกลุ่มโดยคำนวณจากการนำมูลค่าของกลุ่ม APG ใหม่ ($V(S_{apg_k})$) ลบกับมูลค่ากลุ่มของกลุ่ม APG เก่า ($V(S_{apg_j})$) (ในบรรทัดที่ 6) ยกตัวอย่างเช่นพิจารณากลุ่ม 3 ประเทศ คือ {อินโดนีเซีย, มาเลเซีย, บรูไน} เริ่มจากประเทศสมาชิกแรก คือ ประเทศอินโดนีเซียมีมูลค่าของทรัพยากรน้ำมันเป็น 3, ซึ่งเป็นมูลค่าที่เกิดจากการร่วมมือ ล้านบาร์เรล 750 เพียงลำพัง ต่อมาประเทศสมาชิกที่สองคือประเทศมาเลเซีย ทำให้เกิดการเพิ่มมูลค่า โดยการนำมูลค่าของกลุ่มใหม่มาลบกับมูลค่าของกลุ่มเก่าดังนี้ $V(S_{อินโดนีเซีย, มาเลเซีย}) - V(S_{อินโดนีเซีย}) = 9,107 - 3,750 = 5,357$ ซึ่งเป็นมูลค่าที่เกิดจากการร่วมมือของ {อินโดนีเซีย, มาเลเซีย} จากนั้นประเทศที่สามที่เข้ามาเป็นสมาชิกคือประเทศบรูไน ทำให้เกิดการเพิ่มมูลค่า โดยการนำมูลค่ากลุ่มใหม่ลบกับมูลค่ากลุ่มเก่า ดังนี้ $V(S_{อินโดนีเซีย, มาเลเซีย, บรูไน}) - V(S_{อินโดนีเซีย, มาเลเซีย}) = 10,307 - 5,357 = 4,957$ ค่าที่เกิดจากการร่วมมือกันของ {อินโดนีเซีย, มาเลเซีย, บรูไน} และเมื่อพิจารณาการร่วมมือของ {บรูไน, มาเลเซีย, อินโดนีเซีย} การเพิ่มมูลค่าของแต่ละประเทศ เพื่อการร่วมมือกัน ดังนี้

ประเทศบรูไน: $V(S_{บรูไน}) - V(\emptyset) = 1,200 - 0 = 1,200,$ (3)

ประเทศมาเลเซีย: $V(S_{บรูไน, มาเลเซีย}) - V(S_{บรูไน}) = 5,357 - 1,200 = 4,157,$ (4)

$= 6,557 - 1,200 = 5,357$

ประเทศอินโดนีเซีย: (5)

$V(S_{บรูไน, มาเลเซีย, อินโดนีเซีย}) - V(S_{บรูไน, มาเลเซีย}) = 10,307 - 5,357 = 4,957$

และทำซ้ำในกระบวนการคำนวณมูลค่าที่แต่ละประเทศในภูมิภาคอาเซียนเพาเวอร์กริด โดยแบ่งตามแหล่งทรัพยากรที่แต่ละประเทศนำมาเข้ามาเพิ่มยังกลุ่มจนครบ ดังแสดงตัวอย่างมูลค่าที่ได้ในตารางที่ 5

ตารางที่ 5 : แสดงตัวอย่างการจัดสรรส่วนแบ่งของอาเซียนเพาเวอร์กริด

Order	Contribution of APG		
	อินโดนีเซีย	มาเลเซีย	บรูไน
{อินโดนีเซีย,มาเลเซีย,บรูไน}	3,750	5,357	1,200
{อินโดนีเซีย,บรูไน,มาเลเซีย}	3,750	1,200	9,107
{มาเลเซีย,อินโดนีเซีย,บรูไน}	5,357	3,750	6,557
{มาเลเซีย,บรูไน,อินโดนีเซีย}	5,357	1,200	9,107
{บรูไน,อินโดนีเซีย,มาเลเซีย}	1,200	3,750	6,557
{บรูไน,มาเลเซีย,อินโดนีเซีย}	1,200	5,357	4,957
Sum of Contribution	20,614	20,614	37,485
Average of Contribution	3,436	3,436	6,248

จากนั้นทำการหาผลรวมของมูลค่าที่แต่ละเอเจนต์นำมามอบให้กับกลุ่ม (ในบรรทัดที่ 10) ดังตัวอย่างต่อไปนี้

Sum of Contribution (ประเทศอินโดนีเซีย)
 $= 3,750 + 3,750 + 5,357 + 5,357 + 1,200 + 1,200 = 20,614$ (8)

Sum of Contribution (ประเทศมาเลเซีย)
 $= 5,357 + 1,200 + 3,750 + 1,200 + 3,750 + 5,357 = 20,614$ (9)

Sum of Contribution (ประเทศบรูไน)
 $= 1,200 + 9,107 + 6,557 + 9,107 + 6,557 + 4,957 = 37,485$ (10)

จากนั้นจึงทำการหาการกระจายค่าเฉลี่ยของแต่ละประเทศ ใน 6 รายการที่เกิดจากความร่วมมือกัน (12 ในบรรทัดที่) สำหรับค่าเฉลี่ยที่ได้ นั่นคือ

$$\begin{aligned} & 1/6 ((20,614), (20,614), (37,485)) & (11) \\ & = ((3,436), (3,436), (6,248)) \end{aligned}$$

((3,436), (3,436), (6,248)); อินโดนีเซีย, มาเลเซีย, บรูไน) เป็นค่าของเฉลี่ยที่มีการจัดสรรการแบ่งปันส่วนแบ่งแก่ประเทศที่เป็นสมาชิกของอาเซียนเพาเวอร์กริดแต่ละประเทศที่มีทรัพยากรน้ำมันตามลำดับความสำคัญของมูลค่าที่ประเทศนั้นๆนำมามอบให้กับกลุ่มอาเซียนเพาเวอร์กริด

3.4 ผลการทดสอบและการวิจารณ์ผล

ในงานวิจัยนี้ต้องการศึกษาส่วนแบ่งของอาเซียนเพาเวอร์กริด ซึ่งพิจารณาจาก OCS ของอาเซียนเพาเวอร์กริด สามารถกระจายส่วนแบ่งให้กับประเทศในอาเซียนเพาเวอร์กริดที่มีแหล่งทรัพยากรได้อย่างยุติธรรม แบ่งออกเป็นกลุ่มที่มีขนาดเล็ก ขนาดกลาง ขนาดใหญ่ แสดงดังตารางที่ 6

ตารางที่ 6 : แสดงความสัมพันธ์ระหว่างมูลค่ากลุ่มกับขนาดของกลุ่มอาเซียนเพาเวอร์กริด

แหล่งทรัพยากรในอาเซียนเพาเวอร์กริด	ขนาดของกลุ่ม		
	กลุ่มขนาดเล็ก	กลุ่มขนาดกลาง	กลุ่มขนาดใหญ่
แหล่งทรัพยากรน้ำมัน	10.95	4.50	2.03
แหล่งทรัพยากรก๊าซธรรมชาติ	5.31	2.50	8.40
แหล่งทรัพยากรถ่านหิน	10.48	11.25	3.12
แหล่งทรัพยากรพลังน้ำ	5.12	5.20	8.67

จากตารางที่ 6 แสดงผลลัพธ์ในการทดลองลักษณะการกระจายของมูลค่ากับขนาดของกลุ่มอาเซียนเพาเวอร์กริดที่มีแหล่งทรัพยากรน้ำมันปรากฏลักษณะการเกิดส่วนแบ่งในแต่ละ Coalition ของ OCS ที่สามารถแบ่งเป็นกลุ่มขนาดเล็กเป็น 10.95%, กลุ่มขนาดกลาง 4.50%, กลุ่มขนาดใหญ่ 2.03% แหล่งทรัพยากรก๊าซธรรมชาติปรากฏลักษณะการเกิดส่วนแบ่งในแต่ละ Coalition ของ OCS ที่สามารถแบ่งเป็นกลุ่มขนาดเล็ก 5.31%, กลุ่มขนาดกลาง 2.50%, กลุ่มขนาดใหญ่ 8.40% แหล่งทรัพยากรถ่านหินปรากฏลักษณะการเกิดส่วนแบ่งในแต่ละ Coalition ของ OCS ที่สามารถแบ่งเป็นกลุ่มขนาดเล็ก 10.48%, กลุ่มขนาดกลาง 11.25%, กลุ่มขนาดใหญ่ 3.12% และทรัพยากรพลังน้ำปรากฏลักษณะการเกิดส่วนแบ่งในแต่ละ Coalition ของ

OCS ที่สามารถแบ่งเป็นกลุ่มขนาดเล็ก 5.12%, กลุ่มขนาดกลาง 5.20%, กลุ่มขนาดใหญ่ 8.67%

กิตติกรรมประกาศ

ทางผู้วิจัยขอขอบพระคุณคณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคามที่ให้การสนับสนุนทุนในการวิจัย

เอกสารอ้างอิง

- [1] N. R. Suri and Y. Narahari, "Determining the Top-k Nodes in Social Networks using the Shapley Value," in *Autonomous Agents and Multiagent Systems(AAMAS 2008)*, Portugal, 2008, pp. 1509-1512.
- [2] J. M. Bradshaw, *Software Agents*: AAAI Press, 1997.
- [3] V. J. Botti. (2004, Cited January 2004). *Multi-Agent System Technology in a Port Container Terminal Automation* [Online]. Available: http://www.ercim.eu/publication/Ercim_News/enw56/botti.html
- [4] J. e. M. Vidal. (2007). *Fundamentals of Multiagent Systems*. Available: <http://www.damas.ift.ulaval.ca/~coursMAS/Complements/H10/mas-Vidal.pdf>
- [5] M. Wooldridge, *An Introduction to Multiagent Systems*. England: John Wiley & Sons Ltd, 2002.
- [6] M. Davis and M. Maschler, "The Kernel of the Cooperative Game," *Naval Research Logistics Quarterly*, 1965.
- [7] D. B. Gillies, "The Core of a Cooperative Game," in *Springer-Verlag Berlin Heidelberg 2010*, London New York, 2010.
- [8] W. Ketter, J. Collins, P. Reddy, and C. Flath, *The Power Trading Agent Competition*: Erasmus Universiteit Rotterdam, 2011.
- [9] การไฟฟ้าผลิตแห่งประเทศไทย. (2012, Cited 28 ธันวาคม 2556). *ASEAN Power Grid* [Online]. Available: http://www2.egat.co.th/apg/index.php?option=com_content&view=featured&Itemid=435

การพัฒนาการค้นคืนเพลงบนระบบปฏิบัติการ iOS

The Development of Music Information Retrieval System on iOS

กิตติพันธ์ ขันติไตรรัตน์ (Kittipun Khantitirirat), ชยพล ประไพพรเลิศ (Chaiyapol Prapaipornlert)

และ กิติ์สุชาติ พสุภา (Kitsuchart Pasupa)

คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง กรุงเทพฯ

boot191@gmail.com, chainseika@gmail.com, kitsuchart@it.kmitl.ac.th

บทคัดย่อ

บทความนี้นำเสนอการพัฒนาการค้นคืนเสียงเพลง ซึ่งเป็นส่วนหนึ่งของเทคโนโลยีการค้นคืนข้อมูล โดยระบบทำการสกัดลักษณะเด่นของเพลงออกมาเป็นลายพิมพ์เสียง ซึ่งสามารถใช้ในการบ่งบอกถึงลักษณะเฉพาะของแต่ละเพลงได้ ดังนั้นลายพิมพ์เสียงสามารถนำมาใช้ในการระบุเพลงแต่ละเพลงได้ ในปัจจุบันมีผู้พัฒนาแอปพลิเคชันหลายรายได้ทำการพัฒนาระบบค้นหาขึ้นมา อย่างไรก็ตามแอปพลิเคชันเหล่านี้ไม่สามารถค้นหาเพลงไทยได้อย่างสมบูรณ์แบบเมื่อเทียบกับเพลงภาษาอังกฤษ เนื่องจากฐานข้อมูลที่เกี่ยวข้องกับเพลงไทยสากลมีน้อยและไม่สนับสนุนภาษาไทย เราจึงได้พัฒนาระบบการค้นคืนข้อมูลเพลงไทยสากลขึ้นมาเพื่อแก้ไขปัญหาดังกล่าว โดยระบบการค้นคืนเพลงนั้น ได้มีการนำระบบที่มีชื่อว่า 'Echoprint' มาใช้ในการทำงานส่วนหลัง และพัฒนาแอปพลิเคชันบนระบบปฏิบัติการ iOS เป็นระบบส่วนหน้า หลังจากนั้นจึงทำการทดสอบระบบเพื่อหาประสิทธิภาพในการค้นหา

คำสำคัญ: การค้นคืนข้อมูล; การค้นคืนเพลง; ลายพิมพ์เสียง

Abstract

In this paper, we propose the development of music information retrieval system which is one of the information retrieval technologies. The system will extract audio features from an audio sample. We consider this set of features as an audio fingerprint. The audio fingerprint represents the uniqueness of each song. Therefore, this can be used to identify a song. Currently, there are a number of applications which are developed for this purpose. However, their performance for Thai

song retrieval are not good as English ones due to a small number of Thai songs in their database. Moreover, their system are not fully support Thai language. Therefore, we extended the used of an open source music information retrieval system the so-called 'Echoprint' for the system back-end of our system which are fully supported for Thai songs together with an application on iOS. We also compare the performance with existing systems.

Keywords: Information Retrieval, Music Retrieval, Audio Fingerprint.

1. บทนำ

ในอดีตการค้นคืนข้อมูลนั้นสามารถค้นหาได้จากคำสำคัญ (Keyword) ซึ่งระบบทำการค้นหาจากข้อมูลที่มีเมทาดาดา (Metadata) หรือที่ถูกแท็ก (Tag) ไว้ในการค้นคืนข้อมูลเพลง ในสมัยก่อนนั้นเป็นไปได้ยาก เนื่องจากข้อมูลเพลงส่วนมากไม่มีการแท็ก นอกจากชื่อไฟล์ที่มีการระบุชื่อของศิลปินและชื่อเพลงเป็นส่วนใหญ่ และยังไม่มียุทธศาสตร์สืบค้นที่มีประสิทธิภาพเพียงพอกับความต้องการของผู้ใช้

ปัจจุบันระบบการสืบค้นคืนเพลงนั้นได้ถูกพัฒนาให้ดียิ่งขึ้น ระบบสามารถค้นหาข้อมูลเพลงจากเสียงได้ โดยเพลงแต่ละเพลงนั้นถูกอธิบายด้วยลายพิมพ์เสียง (Audio Fingerprint) [1] ซึ่งเป็นลักษณะเด่นที่ได้ถูกสกัดมาจากไฟล์เสียงเพลง และระบบจะทำการใช้ลายพิมพ์เสียงของเสียงเพลงที่ต้องการค้นหา มาเปรียบเทียบกับลายพิมพ์เสียงที่มีอยู่ในฐานข้อมูล โดยผู้ให้บริการสืบค้นข้อมูลเสียงเพลงบนอินเทอร์เน็ตที่นำเทคโนโลยีนี้มาใช้ในปัจจุบันเริ่มมีจำนวนมากขึ้น เช่น โปรแกรมประยุกต์ Shazam [2] และ SoundHound [3] ที่เน้น

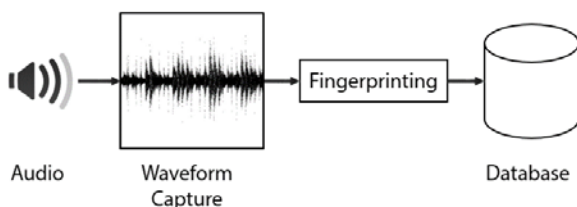
ให้ผู้ใช้งานสามารถค้นหารายละเอียดของเพลงที่กำลังฟังอยู่ รวมถึงการเชื่อมต่อไปยังบริการซื้อเพลงบนมือถือได้ แต่ในปัจจุบัน ระบบนั้นมีปัญหาเมื่อผู้ใช้ต้องการสืบค้นเพลงไทยสากล เนื่องจากฐานข้อมูลของระบบดังกล่าวมีเพลงไทยสากลอยู่จำนวนน้อย ทำให้ไม่สามารถค้นหารายละเอียดได้ และยากแก่การเพิ่มข้อมูลเพลงไทยสากลใหม่ในฐานข้อมูล โดยส่วนของเพลงต่างชาตินั้นสามารถหาได้อย่างไม่มีปัญหาใดๆ อีกทั้งระบบไม่รองรับภาษาไทย ดังนั้นทางผู้จัดทำจึงแก้ปัญหาด้วยการสร้างระบบการสืบค้นเพลงพร้อมทั้งฐานข้อมูลที่เก็บแต่ข้อมูลเพลงไทยสากลอย่างเดียวเท่านั้น โดยพัฒนาต่อออกจากระบบ 'Echoprint' [4]

เนื้อหาของบทความส่วนที่ 2 ได้กล่าวถึงวรรณกรรมที่เกี่ยวข้องกับการวิจัย และพื้นฐานของระบบสืบค้นเพลง ส่วนที่ 3 กล่าวถึงโครงสร้างของแอปพลิเคชันและส่วนต่อประสานกับผู้ใช้ (User Interface) รวมถึงการทดลองสมรรถนะของระบบที่รู้จำเสียงเพลง พารามิเตอร์ที่ใช้และผลที่ได้จากการใช้ระบบ และส่วนที่ 4 เป็นบทสรุปการวิจัยและข้อเสนอแนะ

2. งานวิจัยที่เกี่ยวข้อง

2.1 ลายพิมพ์เสียง

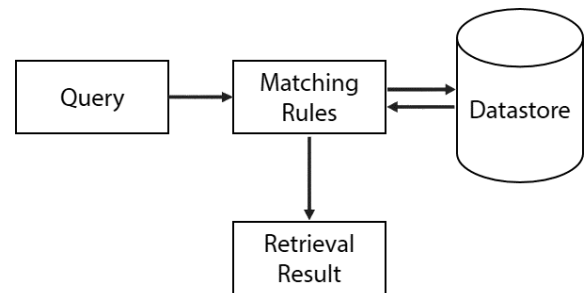
ลายพิมพ์เสียง คือชุดข้อมูลที่บ่งบอกถึงลักษณะเด่นเฉพาะของข้อมูลเสียง สามารถนำมาใช้ในการระบุไฟล์เสียงแต่ละไฟล์ได้ เปรียบได้กับลายนิ้วมือของมนุษย์ที่สามารถบ่งบอกตัวตนได้ ขั้นตอนในการสร้างลายพิมพ์เสียงเริ่มจากการอัดเสียงมาแปลงสัญญาณจากแอนะล็อกเป็นดิจิทัล และใช้เทคนิคต่างๆ ในการสกัดลักษณะเด่น (Feature) จากสัญญาณ เช่น การใช้ Loudness, Brightness, Pitch [1], Power Spectrum [5], Spectrogram [2], Chord Progression [6] เป็นต้น และจัดเก็บไว้ในฐานข้อมูล ดังภาพที่ 1



ภาพที่ 1: กลไกการทำงานของการสร้างลายพิมพ์เสียง

2.2 การสืบค้นข้อมูล (Information Retrieval)

การสืบค้นข้อมูลเป็นเทคโนโลยีที่ใช้ในการค้นหาข้อมูลสารสนเทศต่างๆ โดยระบบการสืบค้นข้อมูลนั้นทำงานโดยการป้อนข้อมูลที่ต้องการค้นหา (Query) เข้าสู่ระบบ และระบบทำการเปรียบเทียบข้อมูลที่ค้นหากับข้อมูลในฐานข้อมูล โดยการใช้เทคนิคในการค้นหาต่างๆ เช่น Term-Weighting [7], Inverted Index [8] เป็นต้น และระบบทำการคืนรายการผลลัพธ์ของการค้นหาค้นกลับมาให้กับผู้ใช้ ดังภาพที่ 2



ภาพที่ 2: หลักการทำงานของ การสืบค้นข้อมูล

โดยเทคนิคการสืบค้นข้อมูลนั้น สามารถใช้งานได้กับข้อมูลหลายรูปแบบ เช่น ข้อความ รูป เสียง เป็นต้น ในบทความนี้จะเน้นเฉพาะการสืบค้นข้อมูลเสียง (Music Information Retrieval) [9] เป็นหลัก โดยสามารถเก็บลายพิมพ์เสียงลงไปในระบบฐานข้อมูล และใช้ในการสืบค้นข้อมูลออกมาได้

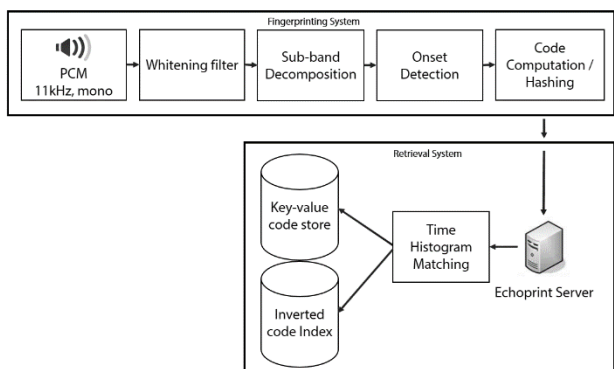
2.3 Echoprint

Echoprint เป็นระบบที่ช่วยในการสืบค้นเพลง โดยทางผู้พัฒนาได้มีการเปิด Source Code ให้ผู้อื่นสามารถสร้างระบบการสืบค้นเพลงได้ด้วยตนเอง ซึ่งส่วนประกอบหลักนั้นมี 2 ส่วน โดยส่วนที่หนึ่งคือระบบสร้างลายพิมพ์เสียงที่ทาง The Echo Nest ได้พัฒนาขึ้นมาที่เรียกว่า 'Codegen' [4] ส่วนที่สองคือส่วนของเซิร์ฟเวอร์ (Server) ซึ่งเป็นส่วนที่ทำหน้าที่สืบค้นข้อมูล ประกอบด้วยระบบ Apache Solr ที่ใช้เป็นระบบการค้นหาและ Tokyo Cabinet เป็นระบบฐานข้อมูลของการค้นหา โดยหลักการของระบบทั้ง 2 ส่วนจะกล่าวถึงในหัวข้อย่อยถัดไป

2.3.1 การสร้างลายพิมพ์เสียงของระบบ

การทำงานของระบบ Codegen นั้นเริ่มจากนำคลื่นเสียงป้อนเข้าไปในระบบ โดยระบบทำงานด้วยการเก็บข้อมูลเป็นคู่ของ {เวลา, แสง} จากสัญญาณเสียงผ่านระบบประมวลสัญญาณและรับมือกับคลื่นรบกวนและการดัดแปลงเสียงต่างๆ โดยแสง (Hash) ในที่นี้หมายถึงการที่นำค่าลักษณะเด่นของช่วงเวลานั้นๆ มาเข้ารหัสและเก็บเป็นข้อมูลที่มีลักษณะเฉพาะโดยวิธีการทำงานแบบละเอียดนั้นเป็นดังนี้

การสร้างลายพิมพ์เสียงนั้นเริ่มจากแปลงคลื่นเสียงที่เข้ามาในระบบให้เป็นคลื่นสัญญาณ Mono ความถี่ 11 kHz หลังจากนั้น สัญญาณจะถูกส่งผ่านไปยัง Whitening Filter [10] เพื่อลดสัญญาณ White Noise ซึ่งเป็นสัญญาณรบกวนที่มีความหนาแน่นคงที่และใช้วิธีการ Sub-band Decomposition [11] ขนาด 8 Band ในการบีบอัดข้อมูลเสียง โดยสัญญาณที่ผ่านจากการทำ Decomposition นั้นจะนำไปใช้ในการหาค้นหา Onset [12] ซึ่งเป็นค่าความเข้มของโน้ต (Note) ในไฟล์เสียง และนำค่า Onset นั้นไปทำการเข้ารหัสขนาด 20 บิต แล้วเก็บคู่กับช่วงเวลาของค่า Onset นั้นๆ หลังจากนั้นทำการนำข้อมูลนี้ส่งไปให้ส่วนของเซิร์ฟเวอร์ การสร้างลายพิมพ์เสียงได้ถูกแสดงอยู่ในภาพที่ 3



ภาพที่ 3: ภาพรวมการทำงานของระบบ Echoprint [4]

2.3.2 การค้นคืนข้อมูลของระบบ

การทำการค้นคืนข้อมูลนั้นมีการทำงานในส่วนของเซิร์ฟเวอร์ โดยระบบจะทำการสร้างดัชนี (Index) ของแต่ละ Onset ในรูปแบบของดัชนีย้อนกลับ (Inverted Index) เก็บไว้ในฐานข้อมูลเพื่อให้สามารถค้นหาได้อย่างรวดเร็ว [8]

ในการเปรียบเทียบนั้น สามารถทำได้โดยการเปรียบเทียบค่าแสงของไฟล์เสียงที่จะค้นหากับค่าแสงในระบบ โดยมีการใช้ดัชนีย้อนกลับเป็นตัวบ่งบอกดัชนี ดังตัวอย่างในตารางที่ 1 ค่าแสงในแถวที่ 1 นั้น มีเพลงที่มีค่าแสงใกล้เคียงกันที่เลขดัชนี 1 และ 7 หลังจากนั้นระบบจะส่งผลลัพธ์การค้นหาเพลงที่ใกล้เคียงที่สุดกลับมายังผู้ใช้งาน แต่อย่างไรก็ตาม ผลลัพธ์ที่ได้จากการใช้ตารางดัชนีย้อนกลับนั้นก็จะพบว่า มีผลลัพธ์ที่ใกล้เคียงกันหลายเพลง ระบบจึงคำนวณ Time Histogram [5] และนำมาใช้ในการคัดเลือกอีกครั้งเพื่อให้ได้ผลลัพธ์ของเพลงที่มีความคล้ายมากที่สุด

ตารางที่ 1: ตารางดัชนีย้อนกลับของค่าแสงที่สกัดออกมา

ข้อมูลที่นำเข้า	เลขดัชนีของเพลง
'eJyIV22SZCkOuxK'	1,7
'Af8xORv8ujNpD2g'	2
'Bb7r7B5985ifhUh'	3,9

3. รายละเอียดการพัฒนา

3.1 ความต้องการของระบบ

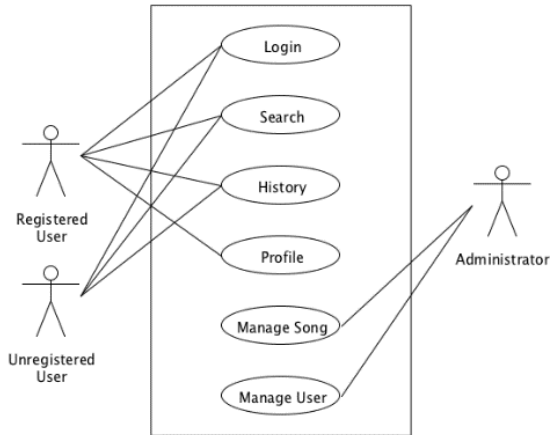
ระบบที่นำเสนอนี้พัฒนามาสำหรับค้นหาเพลงไทยสากลเท่านั้น โดยระบบสามารถวิเคราะห์เสียงเพลงและสามารถค้นคืนรายละเอียดของเพลงไทยสากลที่ป้อนเข้าสู่ระบบได้ เช่น ชื่อเพลง ชื่อนักร้อง ชื่ออัลบั้ม รูปปกอัลบั้มและอื่นๆ พร้อมมีระบบจัดการส่วนของผู้ใช้งาน

โดยผู้ใช้งานที่ยังไม่ได้ลงทะเบียน (Unregistered User) นั้นมีความสามารถในการค้นหาเพลง ดูประวัติการค้นหา เข้าสู่ระบบได้ ส่วนผู้ใช้งานที่ลงทะเบียนกับระบบแล้ว (Registered User) ความสามารถเหมือนกับผู้ใช้ที่ไม่ได้ลงทะเบียน แต่เพิ่มส่วนของประวัติส่วนตัวเข้ามา ส่วนผู้ดูแลระบบ (Administrator) มีความสามารถในการจัดการเพลงในฐานข้อมูล ซึ่งสามารถเพิ่มหรือลบเพลงที่อยู่ในระบบ และจัดการผู้ใช้งานได้ ซึ่งการทำงานของระบบได้ถูกแสดงดังยูสเคสไดอะแกรมในภาพที่ 4

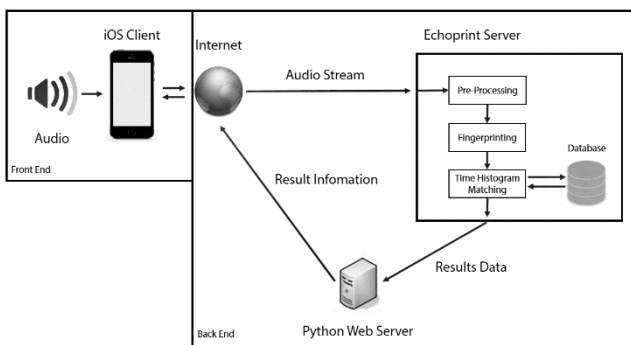
3.2 ภาพรวมของระบบ

ลักษณะการทำงานของระบบนั้นเป็นดังภาพที่ 5 ซึ่งระบบถูกแบ่งออกเป็น 2 ส่วนคือ

3.2.1 ระบบส่วนหน้า (Front End) ถูกพัฒนาอยู่บนระบบปฏิบัติการ iOS เป็นส่วนติดต่อผู้ใช้กับระบบส่วนหลัง โดยผู้ใช้งานสามารถใช้งานผ่านแอปพลิเคชันที่พัฒนาขึ้นมา มีชื่อว่า 'Chordle'



ภาพที่ 4: ยูสเคสไดอะแกรมของระบบ



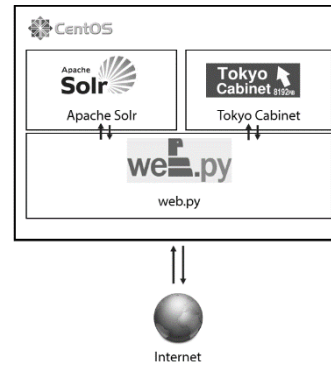
ภาพที่ 5: ภาพรวมการทำงานของระบบ

3.2.2 ระบบส่วนหลัง (Back End) ทำงานอยู่บนระบบปฏิบัติการ CentOS 6.5 [13] ซึ่งแบ่งเป็น 2 ส่วน คือ

3.2.2.1 ระบบค้นคืนเสียงเพลง ซึ่งใช้ระบบ Echoprint ในการสร้างลายพิมพ์เสียงจากไฟล์เสียงที่ได้รับมาเปรียบเทียบกับข้อมูลในฐานข้อมูลและทำหน้าที่ค้นคืนข้อมูล

3.2.2.2 ระบบ Web Server ซึ่งใช้ในการติดต่อกับระบบส่วนหน้า โดยใช้ภาษา Python ในการออกแบบ Web Server

โดยละเอียด ระบบส่วนหลังนั้นมี 3 ระบบที่ใช้ในการทำงานคือ Apache Solr ซึ่งทำหน้าที่เป็นระบบการค้นหา Tokyo Cabinet เป็นระบบฐานข้อมูลและ web.py เป็นระบบเชื่อมต่อระหว่าง Apache Solr และ Tokyo Cabinet รวมถึงรับ-ส่งข้อมูลไปยังแอปพลิเคชัน Chordle ดังภาพที่ 6

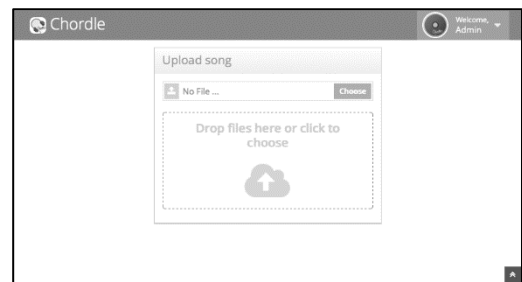


ภาพที่ 6: ภาพรวมการทำงานของระบบส่วนหลัง

3.3 ภาพการทำงานของระบบ

3.3.1 หน้าจอของระบบส่วนหลัง

ผู้ดูแลระบบสามารถเพิ่มเพลงเข้าไปในระบบได้โดยทำการอัปโหลดไฟล์เพลงที่ต้องการ ผ่านหน้าเว็บไซต์ดังในรูปที่ 7



ภาพที่ 7: การเพิ่มเพลงลงในระบบ

3.3.2 หน้าจอของระบบส่วนหน้า

แอปพลิเคชันที่ได้พัฒนาขึ้นมา นั้นมีชื่อว่า Chordle ซึ่งพัฒนาเพื่อระบบปฏิบัติการ iOS 7 เท่านั้น โดยแอปพลิเคชันนี้มีฟังก์ชันการทำงานแบ่งเป็น 3 ส่วน ดังนี้

3.3.2.1 การค้นหา

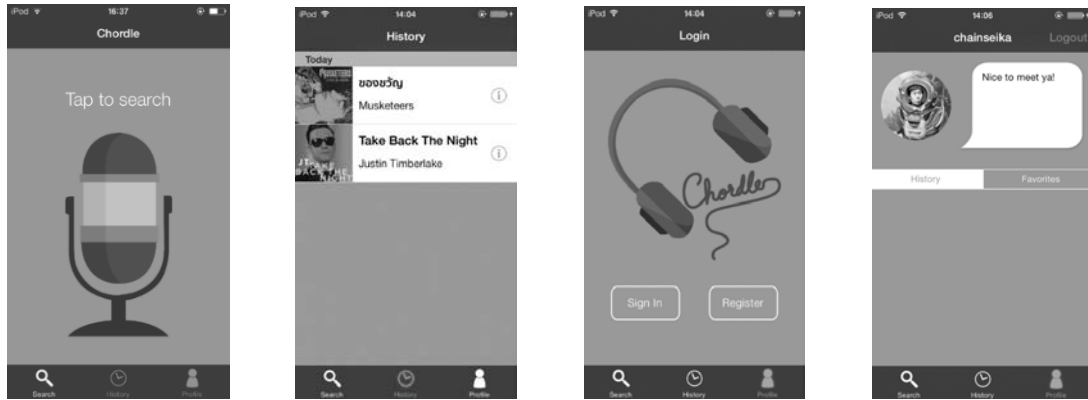
เป็นฟังก์ชันหลักของแอปพลิเคชัน โดยทำหน้าที่ในการค้นหาเพลง ผู้ใช้สามารถใช้งานได้โดยโดยการกดที่รูปไมโครโฟนเพื่อเริ่มต้นในการค้นหา ดังรูปที่ 8(1)

3.3.2.2 ประวัติการค้นหา

ทำหน้าที่ในการแสดงประวัติการค้นหาของผู้ใช้งาน โดยเก็บบันทึกเพลงที่ได้ถูกค้นหาของผู้ใช้งานแอปพลิเคชันในแต่ละวัน ดังรูปที่ 8(2)

3.3.2.3 บัญชีส่วนตัวของผู้ใช้งาน

ทำหน้าที่ในการจัดการบัญชีส่วนตัวของผู้ใช้งาน โดยผู้ใช้งานระบบสามารถลงทะเบียนเพื่อสร้างบัญชีส่วนตัวใน



(1) หน้าจอการค้นหา

(2) หน้าจอประวัติการค้นหา

(3) หน้าจอส่วนก่อน Login

(4) หน้าจอบัญชีส่วนตัว

ภาพที่ 8: รูปฟังก์ชันต่างๆในแอปพลิเคชัน Chordle

ระบบได้ โดยก่อนเข้าสู่ระบบจะมีตัวเลือกให้เลือกว่าจะสร้างบัญชีใหม่หรือทำการเข้าสู่ระบบดังรูปที่ 8(3)

เมื่อทำการเข้าสู่ระบบเรียบร้อยแล้ว จะพบกับหน้าจอข้อมูลของบัญชีที่ผู้ใช้งานสร้างขึ้น ดังรูปที่ 8(4)

4. ผลการทำงานของระบบ

4.1 รูปแบบการทดสอบ

ระบบที่ถูกพัฒนาขึ้นนั้น ได้ถูกทดสอบโดยการค้นหาเพลงไทยสากลทั้งหมด 20 เพลงโดยการเลือกเพลงแบบสุ่มที่มีอยู่ในระบบ โดยแอปพลิเคชัน Chordle จะตั้งเวลาที่ใช้ในการบันทึกเสียงเป็นเวลา 10 15 และ 20 วินาทีตามลำดับ เพื่อทดสอบประสิทธิภาพของระบบ และรายงานผลด้วยเวลาที่ใช้ที่น้อยที่สุดในการบันทึกเสียงในการค้นหาเพลง ส่วนผลการทดลองระบบ Shazam และ SoundHound จะรายงานเป็นหน่วยวินาทีที่ใช้ในการบันทึกเสียงและค้นหาเพลง เนื่องจากไม่สามารถปรับค่าต่างๆ ได้ และทำการหาค่าเฉลี่ยของเวลาที่ใช้ค้นหา อีกทั้งให้ผู้ใช้งานได้ทำการประเมินผ่านแบบสอบถาม

4.2 ปัจจัยที่ใช้ในการทดสอบ

ระบบถูกทดสอบในห้องปฏิบัติการ โดยความดังของเสียงของห้องที่วัดได้มีความดังเฉลี่ยที่ 45 dB อุปกรณ์ที่ใช้ในการทดสอบเป็น iPod Touch 5th Generation และฐานข้อมูลเพลงในระบบนั้นมีทั้งหมด 2,000 เพลง

4.3 ผลการทดสอบ

ตารางที่ 2 แสดงผลการทดลอง โดย Chordle สามารถค้นหาได้ครบ 20 เพลง โดยใช้เวลาเฉลี่ยในการค้นหาที่ 16 วินาที ขณะที่ Shazam สามารถค้นหาได้ทั้ง 18 เพลงจาก

ทั้งหมด 20 เพลง แต่ใช้เวลาเฉลี่ยในการค้นหาที่น้อยที่สุดที่ 4.67 วินาที และ SoundHound สามารถค้นหาได้ 16 เพลงจาก 20 เพลง โดยใช้เวลาเฉลี่ยในการค้นหาที่ 7.94 วินาที

ตารางที่ 2: ผลการทดสอบของระบบ เปรียบเทียบเวลาที่ใช้ในการค้นหารายละเอียดของเพลง โดยที่ – หมายถึงระบบไม่สามารถค้นหาได้

ศิลปิน - ชื่อเพลง	Retrieval Time (s)		
	Chordle	Shazam	SoundHound
25 Hours - คนข้างๆ	15	4	3
แก๊งส้ม - I Love Your Smile	15	5	12
Singular - ล่อง	20	4	6
Stamp - ชายกลาง	15	4	11
Stamp - บ้านเล็ก	15	5	12
Musketeers - นิทาน	15	4	3
Tattoo Colour - รอยจูบ	20	4	12
เบล สุพล - หุดความคิดไม่ได้	15	4	-
นิว & จิว - ไม่รักไม่ต้อง	15	5	5
เบิร์ด ธงไชย - จะได้อะไรคืน	15	5	5
Mild - ดาว	15	-	-
No More Tear - รักเคยดีกว่านี้	20	4	-
ETC - เจ็บและชินไปเอง	15	-	-
The Bottom Blues - ไม่รู้คำว่า	15	5	20
สิงโต นำโชค - อาย	15	6	5
Season Five - แผลรัก	15	3	7
แก๊งส้ม - คุณและคุณเท่านั้น	15	3	4
Bodyslam - แสงสุดท้าย	15	3	11
Lula - ฤดูความรักผลิ	15	4	4
Better Weather - อากาศเปลี่ยนแปลงบ่อย	20	12	7
Average Retrieval Time	16.00	4.20	6.35
Accuracy (%)	100%	90%	80%

จากนั้นได้ทำการเปรียบเทียบคุณสมบัติต่างๆ ของระบบที่นำเสนอและแอปพลิเคชันอื่นๆ ที่มีอยู่ในปัจจุบัน ดังตารางที่ 3

ตารางที่ 3: ตารางเปรียบเทียบคุณสมบัติกับแอปพลิเคชันอื่น

คุณสมบัติ	Chordle	Shazam	SoundHound
การค้นหาเพลงไทย	มี	มี	มี
การอ่านภาษาไทย	รองรับ	ไม่รองรับ	รองรับบางส่วน
ความยืดหยุ่นในการเพิ่มเพลง	มี	ไม่มี	ไม่มี

4.4 ผลการสำรวจความพึงพอใจของผู้ใช้

จากผลการสำรวจความพึงพอใจของผู้ใช้ โดยกลุ่มตัวอย่างคือ นักศึกษาระดับปริญญาตรีจำนวน 20 คน ได้ประเมินความพึงพอใจใน 2 ด้าน คือ ความง่ายของการใช้งานและความสวยงามของแอปพลิเคชัน ซึ่งแบ่งระดับความพึงพอใจเป็น 5 ระดับ ตาม Likert Scale [14] คือ น้อยที่สุด (1), น้อย (2), ปานกลาง (3), มาก (4), และ มากที่สุด (5) ดังตารางที่ 4

ตารางที่ 4: ตารางความพึงพอใจของระบบ

ประเภทความพึงพอใจ	คะแนนเฉลี่ย
ความง่ายต่อการใช้งาน	4.40
ความสวยงาม	4.15
คะแนนเฉลี่ยโดยรวม	4.28

5. สรุปผล

บทความนี้นำเสนอระบบการค้นคืนเพลงบนระบบ iOS จากการศึกษาพบว่าระบบสามารถค้นคืนข้อมูลของเพลงที่มีอยู่ในฐานข้อมูลได้อย่างมีประสิทธิภาพ อย่างไรก็ตามการค้นหายังไม่สามารถเทียบเท่ากับระบบอื่นๆได้ แต่ส่วนระบบสนับสนุนภาษาไทยนั้นสามารถทำงานได้ดี ผู้ใช้ให้ความพึงพอใจต่อแอปพลิเคชันอยู่ในระดับมาก ที่คะแนน 4.28 อีกทั้งระบบยังมีความยืดหยุ่นในการเพิ่มเพลงใหม่อีกด้วย

ในอนาคต สามารถพัฒนาประสิทธิภาพระบบได้โดยการพัฒนาอัลกอริทึมที่ใช้ในการสร้างลายพิมพ์เสียงและการค้นคืน อีกทั้งมีการเพิ่มข้อมูลเพลงเข้าไปในระบบอย่างสม่ำเสมอ

เอกสารอ้างอิง

[1] P. Cano, E. Battle, E. Gomez, L.T. Gomes and M. Bonnet, "Audio Fingerprinting: Concepts and Applications" *Computational Intelligence for Modelling and Prediction*, pp. 233-245, 2005.

[2] A. Wang, "An Industrial-Strength Audio Search Algorithm." *In Proceeding of the 4th International Conference on Music Information Retrieval*, 2003.

[3] SoundHound, "Sound2Sound (S2S) Search Science." [Internet] Available at: <http://www.soundhound.com/index.php?action=s.sound2sound>. [23 Jan 2014]

[4] D.P.W. Ellis, B. Whitman and A. Porter, "Echoprint - An Open Music Identification Service." *In Proceeding of the 12th International Society for Music Information Retrieval Conference*, 2011.

[5] W. Liang, S. Zhang and B. Xu, "A Histogram Algorithm for Fast Audio Retrieval." *In Proceeding of the 6th International Conference on Music Information Retrieval*, pp. 586 - 589, 2005.

[6] B. Absolu, T. Li and M. Ogihara, "Analysis of Chord Progression Data" *Advance in Music Information Retrieval*, Vol. 274, pp. 165-184, 2010.

[7] G. Salton and C. Buckley, "Term-weighting Approaches in Automatic Text Retrieval." *Information Processing & Management*, Vol. 24, No. 5, pp. 513-523, 1988.

[8] A.K. Mahapatra and S. Biswas, "Inverted Indexes: Types and Techniques." *IJCSI International Journal of Computer Science Issues*, Vol. 8, Issue 4, No 1, pp. 384-392, 2011.

[9] J.S. Downie, "Music Information Retrieval." *Annual Review of Information Science and Technology*, Vol. 37, Issue 1, pp. 295-340, 2003.

[10] A. Krishnamoorthy, "Linear Extended Whitening Filters." *Information Theory, Cornell University Library*, 2013.

[11] P. Motlicek, S. Ganapathy, H. Hermansky, H. Garudadri and M. Athineos, "Perceptually Motivated Sub-band Decomposition for FDLP Audio Coding" *In Proceeding of the 11th International Conference on Text, Speech and Dialogue*, pp. 435-442, 2008.

[12] J. P. Bello, L. Daudet, S. Abdallah, C. Duxbury, M. Davies and M.B. Sandler, "A Tutorial on Onset Detection in Music Signals." *IEEE Trans. on Speech and Audio Processing*, Vol. 13, No. 5, pp. 1035 - 1047, 2005.

[13] The CentOS Project, "CentOS Linux." [Internet] Available at: <http://www.centos.org>. [27 Jan 2014]

[14] K. L. Wuensch. What is a Likert Scale? and How Do You Pronounce 'Likert?'. East Carolina University. [Internet] Available at: <http://core.ecu.edu/psyc/wuenschk/StatHelp/Likert.htm>. [10 Jan 2014]

การศึกษาความสัมพันธ์ระหว่างการแสดงออกพฤติกรรมของเด็กนักเรียนระดับมัธยม หลังจากที่เคยถูกรังแกด้วยพฤติกรรมเดียวกันด้วยเทคนิคกฎความสัมพันธ์ High School Student Behavior Study Using Association Rule Technique

ทวีศักดิ์ น้อยญาติ (Taweesak Noipasee)¹ ทเวา คำปาเชื้อ (Tawa Khampachua)² และ จริฎุ แสนราช (Charun Sanrch)³
ภาควิชาคอมพิวเตอร์ศึกษา คณะครุศาสตร์อุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
taweesaknoipasee@gmail.com¹, twk@kmutnb.ac.th², jsr@kmutnb.ac.th³

บทคัดย่อ

บทความนี้นำเสนอผลการศึกษาความสัมพันธ์ของข้อมูลพฤติกรรมของนักเรียนระดับมัธยมที่พวกเขาถูกรังแกและพฤติกรรมที่พวกเขาแสดงออกกับคนรอบตัว ข้อมูลที่นำมาศึกษาเป็นข้อมูลที่ได้จากการตอบแบบสอบถามโดยนักเรียนระดับมัธยม การวิเคราะห์ใช้เทคนิคการค้นหากฎความสัมพันธ์เพื่อค้นหารูปแบบของความสัมพันธ์ระหว่างพฤติกรรมที่นักเรียนระดับมัธยมถูกรังแกโดยคนรอบตัวในบริบทต่างๆ ว่าเกี่ยวข้องกับพฤติกรรมที่พวกเขาแสดงออกหรือกระทำกับคนรอบข้างในบริบทต่างๆ หรือไม่ ซึ่งผลการศึกษาพบว่าเด็กนักเรียนมัธยมจะแสดงออกหรือกระทำพฤติกรรมบางอย่างที่ไม่ดีกับเพื่อนเมื่อพวกเขาถูกรังแกพฤติกรรมบางอย่างที่ไม่ดีจากคนรอบตัวที่มีวิญญูณมากกว่า โดยมีค่าความเชื่อมั่นของความสัมพันธ์มากกว่าร้อยละ 90 คือ ถ้าพวกเขาถูกพ่อแม่ ครู และคนในชุมชนรังแกด้วยการตวาด ตะคอก แผลเสียด พุดประชดประชัน บิดเนื้อหรือบิดหู พวกเขาจะไปแสดงออกหรือกระทำพฤติกรรมการเขกหัว ดบตีแขนหรือหัว พุดล้อเลียนให้รู้สึกลาย และคำด้วยคำหยาบคาย กับเพื่อน

คำสำคัญ: กฎความสัมพันธ์ เทคนิคเหมืองข้อมูล พฤติกรรม ความรุนแรง

Abstract

This paper presents the study of relationship between how the secondary school students are being treated and how they behave and treat the others. By using association rule, one can find whether the way in which the students are being treated has impact on how they treat the others in different situations or not. According

to our study, we found that the secondary school students would express inappropriate behavior or mistreat their friends when they were being mistreated by the others who were senior to them at the minimum confidence value of 90%. More specifically, if they were scolded, yelled, sarcastic, or got pinched at the ears by their parents, teachers, or other people within their community, they would mistreat their friends by hitting on their friends' heads or hands, teasing and embarrassing them, and impolitely curse their friends.

Keyword: Association Rule, Apriori Algorithms, Data Mining, Violent Behavior

1. บทนำ

เมื่อเด็กหรือเยาวชนถูกรังแกด้วยพฤติกรรมบางอย่างจากคนรอบตัวในสังคมแล้ว พวกเขาจะซึมซับเอาพฤติกรรมเหล่านั้นไปแสดงออกหรือกระทำกับคนรอบตัวต่อหรือไม่เป็นประเด็นการวิจัยที่บทความนี้ต้องการศึกษา ที่ผ่านมามีทั้งจากการศึกษาและประสบการณ์ของผู้วิจัยเองพบว่าการเรียนการสอนผ่านการรับรู้นั้นส่งผลต่อการเปลี่ยนแปลงพฤติกรรมของผู้เรียน ซึ่งประสบการณ์การเรียนรู้ของผู้เรียนนั้นเกิดได้หลายสถานที่ทั้งในชั้นเรียนในรูปแบบที่เป็นทางการและนอกชั้นเรียนที่บ้านหรือที่ชุมชนในรูปแบบที่ไม่เป็นทางการ เมื่อผู้เรียนรับรู้และเข้าใจในประสบการณ์ใดแล้วพวกเขาก็สามารถเปลี่ยนแปลงพฤติกรรมของตนเองตามประสบการณ์นั้นได้แต่คนส่วนใหญ่จะมีมุมมองต่อเนื้อหาหรือประสบการณ์ในด้านที่ดีเท่านั้น ถ้าเป็นอย่างนั้นจริงย่อมไม่ดีแน่ถ้าพวกเขาเหล่านั้นรับรู้และเข้าใจในประสบการณ์ผ่านทางพฤติกรรมในด้านที่ไม่ดี เนื่องจากการแสดงออกพฤติกรรมที่ไม่ดีนั้นก่อให้เกิดปัญหา

ของการอยู่ร่วมกันในสังคม และอาจจะพัฒนาไปสู่การกระทำพฤติกรรมที่รุนแรงต่อไป

เพื่อให้สังคมไทยเป็นสังคมที่ผู้คนอยู่ร่วมกันอย่างสันติปราศจากการใช้ความรุนแรงต่อกัน ผู้ที่รับผิดชอบควรมีการศึกษาและวิเคราะห์หาต้นตอของปัญหาเพื่อนำมาสร้างมาตรการหรือแนวทางเชิงรุกที่สามารถยับยั้งหรือป้องกันต่อไป ดังที่ได้กล่าวมาข้างต้นงานวิจัยนี้สนใจประเด็นกระบวนการเรียนรู้ในพฤติกรรมที่ไม่ดีของกลุ่มเด็กและเยาวชนผ่านประสบการณ์ที่พวกเขาถูกกระทำโดยคนรอบตัวนั้นจะส่งผลให้พวกเขาแสดงออกพฤติกรรมนั้นในอนาคตหรือไม่ โดยผู้วิจัยได้ศึกษาและวิเคราะห์ข้อมูลพฤติกรรมของเด็กและเยาวชนในระดับมัธยม โดยนำข้อมูลที่ได้จากการตอบแบบสอบถามแบบออนไลน์ของเด็กมัธยมในเขตกรุงเทพมหานครและจังหวัดนนทบุรีมาทำการวิเคราะห์ด้วยเทคนิคกฎความสัมพันธ์ (Association Rule) [1] เพื่อการหารูปแบบของความสัมพันธ์ว่าพฤติกรรมที่เด็กและเยาวชนระดับมัธยมถูกกระทำในบริบทต่างๆ แต่ละพฤติกรรมนั้นส่งผลต่อการแสดงออกพฤติกรรมกับคนรอบข้างในบริบทต่างๆ ของตัวเขาเหล่านั้นหรือไม่ ผลที่ได้จากการศึกษานี้สามารถนำมาสร้างเป็นข้อสรุปของสาระในฐานข้อมูลเพื่อกำหนดนโยบายและมาตรการในการป้องกันและแก้ไขปัญหาการแสดงออกพฤติกรรมความรุนแรงในเด็กและเยาวชนต่อไป

การนำเสนอบทความแบ่งออกเป็น 4 หัวข้อ แนวคิดและวิธีการทำเหมืองข้อมูลนำเสนอในหัวข้อต่อไป หัวข้อที่ 3 นำเสนองานวิจัยทางด้านพฤติกรรมความรุนแรง หัวข้อที่ 4 ผลที่ได้จากการค้นหาความสัมพันธ์ และหัวข้อสุดท้ายนำเสนอการอภิปรายและสรุปผลที่ได้จากการวิจัย

2. แนวคิดและวิธีการทำเหมืองข้อมูล

การทำเหมืองข้อมูล คือ กระบวนการค้นหาสารสนเทศหรือข้อความรู้ที่อยู่ในฐานข้อมูลขนาดใหญ่ที่ซับซ้อน [2] เพื่อนำข้อความรู้ที่ได้ไปใช้ประโยชน์ในการสร้างการพยากรณ์ ตัวแบบสำหรับการจำแนกหน่วยหรือกลุ่มข้อมูล แสดงความสัมพันธ์ระหว่างหน่วยข้อมูล และสร้างข้อสรุปของสาระในฐานข้อมูล ที่ต้องอาศัยเทคนิคหรือวิธีการ ได้แก่ วิธีการจัดกลุ่ม วิธีการจำแนก การค้นหาความสัมพันธ์ การพยากรณ์ เป็น

ต้น โดยการสร้างตัวแบบที่อธิบายความเป็นไปหรือสภาพการณ์หนึ่งที่เกิดขึ้นแล้ว เพื่อนำตัวแบบนี้มาใช้อธิบายสถานการณ์ที่ยังไม่เกิดขึ้นหรือที่ยังไม่รู้คำตอบ ดังนั้นถ้ามีฐานข้อมูลขนาดใหญ่ที่มีข้อมูลคุณภาพดีเทคโนโลยีการทำเหมืองข้อมูลสามารถช่วยในการค้นหาหรือแสวงหาโอกาสหรือแนวทางใหม่ ก่อให้เกิดกระบวนการอัตโนมัติ การพยากรณ์แนวโน้มและพฤติกรรม และการค้นพบรูปแบบที่ไม่เคยรู้จักมาก่อนที่ซ่อนอยู่ในฐานข้อมูลนั้น

2.1 กฎความสัมพันธ์

การค้นหาความสัมพันธ์เป็นเทคนิคหนึ่งที่สำคัญของการทำเหมืองข้อมูล (Data Mining) หลักการที่สำคัญของเทคนิคนี้ คือ ค้นหาความสัมพันธ์ระหว่างรายการข้อมูลที่ปรากฏร่วมกัน (Co-occurrence) ในฐานข้อมูลขนาดใหญ่ เพื่อนำผลลัพธ์ที่อยู่ในรูปของกฎความสัมพันธ์ไปใช้ในการทำความเข้าใจหรือทำนายปรากฏการณ์ที่เป็นบริบทของฐานข้อมูลนั้นๆ กฎความสัมพันธ์สามารถเขียนในรูปแบบกลุ่มของรายการข้อมูล (Item Set) ที่เป็นเหตุไปสู่กลุ่มของรายการข้อมูลที่เป็นผล รากฐานของเทคนิคนี้มาจากการวิเคราะห์พฤติกรรมการซื้อสินค้า (Market Basket Analysis) ที่นักการตลาดต้องการวิเคราะห์หารายการสินค้าที่มีการซื้อพร้อมกัน ณ จุดขาย โดยผลลัพธ์ที่ได้สามารถแสดงในรูปแบบของกฎ ที่แสดงความเป็นไปได้ของการซื้อสินค้าต่างๆ ร่วมกัน ยกตัวอย่าง เช่น ลูกค้าส่วนใหญ่ที่มาซื้อหนังสือมติดิชนรายสัปดาห์แล้วมักจะซื้อวารสารคู่สร้างคู่สมด้วย ซึ่งทางร้านสามารถนำความรู้นี้ไปแนะนำลูกค้าใหม่ที่กำลังซื้อหนังสือมติดิชนรายสัปดาห์ได้ โดยกฎความสัมพันธ์นี้ใช้ในการหาความสัมพันธ์ของข้อมูลทั้งสองชุด หรือมากกว่าสองชุดขึ้นไปมาไว้ด้วยกัน ระดับความสัมพันธ์ของกฎพิจารณาจากตัวแปรสองตัว คือ ค่าสนับสนุน (Support) ซึ่งเป็นค่าร้อยละของการดำเนินการที่กฎสามารถนำไปใช้ หรือ ค่าร้อยละของการดำเนินการที่กฎที่ใช้มีความถูกต้อง และค่าความมั่นใจ (Confidence) ซึ่งเป็นจำนวนของกรณีที่ถูกถูกต้องโดยสัมพันธ์กับจำนวนของกรณีที่ถูกสามารถนำไปใช้ได้

2.2 การค้นหากฎความสัมพันธ์

ขั้นตอนการค้นหาความสัมพันธ์แบ่งออกเป็น 2 ขั้นตอน คือ การสืบค้นความถี่ของกลุ่มรายการ (Frequent Item

Set) ที่เป็นการพิจารณาความมั่นใจและกลุ่มรายการที่มีความถี่หรือค่าสนับสนุนว่ามีค่ามากกว่าหรือเท่ากับค่าสนับสนุนขั้นต่ำ

การค้นหากฎความสัมพันธ์สามารถยกตัวอย่างได้ดังนี้ ถ้าตารางที่ 1 เป็นตัวอย่างชุดข้อมูลของการซื้อสินค้า โดยคอลัมน์ TID รายการสินค้าที่ถูกซื้อโดยลูกค้าแต่ละคนและคอลัมน์ ITEMS คือรายการสินค้าที่ซื้อพร้อมกันในแต่ละครั้งโดย A, B, C, D และ E แทนซื้อสินค้าในแต่ละชนิด การดำเนินการกำหนดค่าสนับสนุนขั้นต่ำ (Minimum Support) เท่ากับร้อยละ 50 และค่าความมั่นใจขั้นต่ำ (Minimum Confidence) เท่ากับร้อยละ 70 โดยทั่วไปการกำหนดค่าขั้นต่ำของทั้งสองค่าขึ้นอยู่กับความต้องการของผู้ใช้หรือข้อเสนอแนะของผู้เชี่ยวชาญ

ตารางที่ 1: ตัวอย่างข้อมูลรายการซื้อสินค้า

TID	ITEMS
1	A,C,D
2	B,C,E
3	A,B,C,E
4	B,E
5	A,B,C,E

ทำการค้นหา Frequent Itemsets รอบที่ 1 รายการสินค้าที่มีค่าสนับสนุนมากกว่าร้อยละ 50 จะถูกตัดข้อมูลทิ้งและได้ผลลัพธ์ดังตารางที่ 2

ตารางที่ 2: ผลการค้นหา Frequent Itemsets รอบที่ 1

Itemsets C1	Count	%
{A}	3	60
{B}	4	80
{C}	4	80
{D}	1	20
{E}	4	80

การค้นหา Frequent Itemsets รอบที่ 2 ก็ทำลักษณะเดียวกันกับรอบแรกโดยได้ผลลัพธ์ดังตารางที่ 3

ตารางที่ 3: ผลการค้นหา Frequent Itemsets รอบที่ 2

Itemsets C2	Count	%
{A,B}	2	40
{A,C}	3	60
{A,E}	2	40
{B,C}	3	60
{B,E}	4	80
{C,E}	3	60

การค้นหา Frequent Itemsets ทำไปจนถึง Itemset ทั้งหมดโดยได้ผลลัพธ์ดังตารางที่ 3

ตารางที่ 4: ผลการค้นหา Frequent Itemsets รอบสุดท้าย

Itemsets C3	Count	%
{B,C,E}	3	60

กฎความสัมพันธ์ที่ถูกสร้างขึ้นสามารถสรุปได้ดังตารางที่ 5

ตารางที่ 5: รายการกฎความสัมพันธ์ที่ค้นพบ

กฎความสัมพันธ์	ค่าสนับสนุน (%)	ค่าความเชื่อมั่น (%)
{BC}=>{E}	60	3/3=100
{CE}=>{B}	60	3/3=100
{BE}=>{C}	60	3/4=75
{B}=>{CE}	60	3/4=75
{C}=>{BE}	60	3/4=75
{C}=>{BE}	60	3/4=75
{E}=>{BC}	60	3/4=75
{A}=>{C}	60	3/3=100
{B}=>{C}	60	3/4=75
{B}=>{E}	80	3/3=100
{C}=>{E}	60	3/4=75

เมื่อพิจารณาจากกฎความสัมพันธ์พบว่า {BC} => {E} มีค่าสนับสนุนเท่ากับร้อยละ 60 และค่าความเชื่อมั่นเท่ากับ 100% หมายความว่าจำนวนครั้งการซื้อสินค้าที่มีการซื้อ B, C, E พร้อมกันมีจำนวน 3 ครั้งจากจำนวนรายการทั้งหมด และความน่าจะเป็นเมื่อมีการซื้อ B, C พร้อมกันแล้วจะซื้อสินค้า E ด้วยเสมอที่ความเชื่อมั่น 100%

3. การศึกษาและวิจัยด้านพฤติกรรมความรุนแรง

จากการศึกษารายงานการศึกษาวิจัยขององค์กรสหประชาชาติพบว่าเด็กต้องเผชิญกับความรุนแรงทั้งกาย วาจา และเพศจากบุคคลใกล้ชิด เช่น พ่อแม่ เพื่อน ครูอาจารย์ และคนในชุมชน โดยถูกกระทำทั้งที่บ้าน โรงเรียน สถานที่ให้คุ้มครองดูแลเด็ก และชุมชน [3] ซึ่งเป็นปัญหาที่ทั่วโลกให้ความสนใจเป็นอย่างมาก นอกจากนั้น ได้มีการเรียกร้องให้ผู้ที่เกี่ยวข้องของแต่ละประเทศจัดทำนโยบาย ยุทธศาสตร์ และแผนปฏิบัติการเพื่อแก้ไขปัญหาดังกล่าว

ในประเทศไทยหน่วยงานที่เกี่ยวข้องได้ตระหนักถึงความสำคัญของปัญหาความรุนแรงต่อเด็กเช่นเดียวกัน ได้มีการดำเนินการศึกษาเรื่องความรุนแรงต่อเด็ก เพื่อใช้เป็นข้อมูลพื้นฐานในการกำหนดนโยบายและมาตรการเพื่อการป้องกันและแก้ไขปัญหาคความรุนแรงต่อเด็กให้สอดคล้องกับสถานการณ์ที่เกิดขึ้น ให้สังคมตระหนักถึงสถานการณ์ความรุนแรงต่อเด็ก ให้ทราบทัศนคติเกี่ยวกับความรุนแรง และการปฏิบัติต่อกันระหว่างเด็กกับผู้ใหญ่ เพื่อนำไปสู่ความเข้าใจในปัญหา ความร่วมมือในการป้องกัน และการแก้ปัญหา

การศึกษาเกี่ยวกับความรุนแรงต่อเด็กในประเทศไทยเป็นการรวบรวมข้อมูลประสบการณ์เกี่ยวกับความรุนแรงในวัยเด็กที่ได้จากการสัมภาษณ์ ในเบื้องต้นผลของการศึกษาเป็นเพียงรายงานการสรุปในภาพรวมและการให้ข้อคิดเห็นจากการแปรผลทางสถิติเท่านั้นยังไม่ได้มีการวิเคราะห์หาสารสนเทศและข้อความรู้ที่แฝงอยู่ในข้อมูล [3]

งานวิจัยนี้ได้ใช้แนวทางเดียวกับการศึกษาข้างต้น คือผู้วิจัยได้จัดทำแบบสอบถามแบบออนไลน์เพื่อเก็บข้อมูลกับนักเรียนระดับมัธยม โดยรายละเอียดของข้อมูลประกอบด้วยประสบการณ์ความรุนแรงที่เกิดจากการกระทำต่อกันทั้งกาย วาจา และเพศจากรอบข้าง เช่น พ่อแม่ เพื่อน ครูอาจารย์ และคนในชุมชน ทั้งที่บ้าน โรงเรียน และชุมชน โดยแบ่งออกเป็นสองตอน ประกอบด้วยประสบการณ์ความรุนแรงที่เขาถูกกระทำ และพฤติกรรมความรุนแรงที่เขาเป็นผู้กระทำ

4. ผลการค้นหากฎความสัมพันธ์

4.1 วิธีการวิเคราะห์

ข้อมูลที่ได้จากการตอบแบบสอบถามมีจำนวน 4251 คน จากนั้นได้นำมาพิจารณาหาความถี่ของพฤติกรรมที่ถูกกระทำ

เพื่อนำมาเลือกเป็น Item ในการหาความเชื่อมโยงซึ่งได้พิจารณาจากค่าความถี่ที่เกิดขึ้นสูงสุด จำนวน 12 รายการโดยสามารถสรุปรายการของพฤติกรรมได้ดังตารางที่ 6

ตารางที่ 6: รายการพฤติกรรมที่มักเกิดการกระทำจริงโดยคนรอบตัวเด็กนักเรียนมัธยม

Item	รายการพฤติกรรม
I1	พฤติกรรม บิดเนื้อหรือบิดหูกระทำโดยพ่อแม่
I2	ลงโทษด้วยการเฉี้นกระทำโดยพ่อแม่
I3	พฤติกรรมตวาด ตะคอก แผลเสียงกระทำโดยครู
I4	พฤติกรรมพูดประชดประชันที่กระทำโดยครู
I5	พฤติกรรมพูดลามกกระทำโดยครู
I6	พฤติกรรมพูดลามกทางเพศกระทำโดยครู
I7	เราแสดงพฤติกรรมโดยเขหัวเพื่อน
I8	เราแสดงพฤติกรรมโดยตบหรือตีแขนหรือหัวเพื่อน
I9	เราแสดงพฤติกรรมโดยล้อเลียนเพื่อนให้รู้สึกอาย
I10	เราแสดงพฤติกรรมโดยด่าว่าเพื่อนด้วยคำหยาบคาย
I11	เราแสดงพฤติกรรมโดยพูดลามกกับเพื่อน
I12	เราให้ดูรูปภาพ วิดีโอ หรือคลิปโป๊

การวิเคราะห์ใช้แนวคิดการค้นหากฎความสัมพันธ์ หรือ อพริออริอัลกอริทึม (Apriori Algorithms) [4][5] ซึ่งเป็นอัลกอริทึมในการสืบค้นกฎความสัมพันธ์ที่เกิดขึ้นได้ทั้งหมด โดยพิจารณาจากค่าสนับสนุนขั้นต่ำที่กำหนด ผลลัพธ์ในการสืบค้นจะอยู่ในรูปแบบ Lattice ที่แสดง ถึงความสัมพันธ์ของ Itemset หรือกลุ่มของรายการข้อมูลที่ปรากฏในฐานข้อมูล โดยมีการกำหนดค่าสนับสนุนขั้นต่ำสุดที่ 0.1 หรือร้อยละ 10 เหตุผลที่เลือกเป็นร้อยละ 10 เนื่องจากค่าทางสถิติที่ให้ความถูกต้องนั้นตามทฤษฎีจะใช้ขั้นต่ำที่ร้อยละ 10 ค่าความเชื่อมั่นกำหนดที่ 0.9 หรือร้อยละ 90 เพื่อให้ได้กฎความสัมพันธ์ที่มีค่าความถูกต้องร้อยละ 90 ขึ้นไป และกำหนดจำนวนกฎเพื่อหยุดการคำนวณที่ $2^{13}-1$ หรือ 8191 กฎ

4.2 ผลการค้นหากฎความสัมพันธ์

การค้นหากฎความสัมพันธ์ด้วยอพริออริอัลกอริทึม ที่ค่าความเชื่อมั่นร้อยละ 90 ขึ้น ไปจำนวน 6 กฎด้วยกัน ดังนี้

กฎที่ 1 ค่าความเชื่อมั่นที่ร้อยละ 96

ถ้าครูพูดแบบนี้กับเรา ตวาด ตะคอก แผลเสียด พุดประชดประชัน เราไปเขกหัวเพื่อน ตบตีแขนหรือหัวเพื่อน พุดล้อเรียนเพื่อนให้รู้สึกอาย พฤติกรรมต่อไปเราจะค่าเพื่อนด้วยคำหยาบคาย

กฎที่ 2 ค่าความเชื่อมั่นที่ร้อยละ 96

ถ้าโรงเรียนทำแบบนี้กับเรา บิดเนื้อหรือบิดหู ครูพูด ตวาด ตะคอก แผลเสียด กับเรา เราไปเขกหัวเพื่อน ตบตีแขนหรือหัวเพื่อน พุดล้อเรียนเพื่อนให้รู้สึกอาย พฤติกรรมต่อไปเราจะค่าเพื่อนด้วยคำหยาบคาย

กฎที่ 3 ค่าความเชื่อมั่นที่ร้อยละ 95

ถ้านักเรียนเป็นนักเรียนระดับมัธยมศึกษาตอนปลาย ครูพูดแบบนี้กับเรา ตวาด ตะคอก แผลเสียด เราไปเขกหัวเพื่อน ตบตีแขนหรือหัวเพื่อน ค่าเพื่อนด้วยคำหยาบคาย พฤติกรรมต่อไปเราจะพุดล้อเรียนเพื่อนให้รู้สึกอาย

กฎที่ 4 ความเชื่อมั่นที่ร้อยละ 95

ถ้านักเรียนเป็นนักเรียนระดับมัธยมศึกษาตอนปลาย ถ้าครูพูดแบบนี้กับเรา ตวาด ตะคอก แผลเสียด เราไปเขกหัวเพื่อน ตบตีแขนหรือหัวเพื่อน พุดล้อเรียนเพื่อนให้รู้สึกอาย พฤติกรรมต่อไปเราจะค่าเพื่อนด้วยคำหยาบคาย

กฎที่ 5 ความเชื่อมั่นที่ร้อยละ 94

ถ้าโรงเรียนทำแบบนี้กับเรา บิดเนื้อหรือบิดหู ครูพูด ตวาด ตะคอก แผลเสียด กับเรา เราไปเขกหัวเพื่อน ตบตีแขนหรือหัวเพื่อน เราพุดกับเพื่อน ค่าเพื่อนด้วยคำหยาบคาย พฤติกรรมต่อไปเราจะพุดล้อเรียนเพื่อนให้รู้สึกอาย

กฎที่ 6 ความเชื่อมั่นที่ร้อยละ 94

ถ้าครูพูดแบบนี้กับเรา ตวาด ตะคอก แผลเสียด พุดประชดประชัน เราไปเขกหัวเพื่อน ตบตีแขนหรือหัวเพื่อน เราพุดกับเพื่อน ค่าเพื่อนด้วยคำหยาบคาย พฤติกรรมต่อไปเราจะพุดล้อเรียนเพื่อนให้รู้สึกอาย

5. อภิปรายและสรุปผลที่ได้จากการวิจัย

การศึกษาความสัมพันธ์ของข้อมูลพฤติกรรมที่ถูกกระทำและพฤติกรรมที่แสดงออกกับคนรอบตัวของนักเรียนระดับมัธยมที่ตอบแบบสอบถามแบบออนไลน์จำนวน 4251 คนโดยใช้เทคนิคการค้นหากฎความสัมพันธ์พบว่าเด็กนักเรียนมัธยมจะแสดงออกหรือกระทำพฤติกรรมบางอย่างที่ไม่ดีกับเพื่อนเมื่อพวกเขาถูกกระทำพฤติกรรมบางอย่างที่ไม่ดีจากคนรอบตัวที่มีวิญญูมากกว่า โดยมีค่าความเชื่อมั่นของความสัมพันธ์มากกว่าร้อยละ 90 คือ ถ้าพวกเขาถูกพ่อแม่ ครู และคนในชุมชนกระทำด้วยการตวาด ตะคอก แผลเสียด พุดประชดประชัน บิดเนื้อหรือบิดหู พวกเขาจะไปแสดงออกหรือกระทำพฤติกรรมการเขกหัว ตบตีแขนหรือหัว พุดล้อเรียนให้รู้สึกอาย และค่าด้วยคำหยาบคาย กับเพื่อน

ผลการที่ได้จากการศึกษาข้อมูลนี้สามารถตั้งข้อสังเกตได้ว่าการแสดงออกพฤติกรรมที่ไม่เหมาะสมบางอย่างของเด็กและเยาวชนนั้นมีอิทธิพลมาจากการที่เขาถูกกระทำโดยคนที่มีความวิญญูมากกว่า โดยพวกเขาจะไม่มีโอกาสโต้ตอบแต่จะแสดงพฤติกรรมในลักษณะเดียวกันกับคนในวิญญูระดับเดียวกัน

เอกสารอ้างอิง

- [1] กฤษณะ ไวยมัย , ชิดชนก ส่งศิริ และ ธนาวิทย์ รักธรรมานนท์, การใช้เทคนิคดาต้า ไมน์นิงเพื่อพัฒนาคุณภาพการศึกษานิสิตคณะวิศวกรรมศาสตร์ , *NECTEC Technical Journal*, Vol 3, No. 11, July-Oct, 2001.
- [2] สุชาติ ภิระนันท์, คำถามคำตอบเกี่ยวกับการทำเหมืองข้อมูลเบื้องต้น, เอกสารประกอบการสัมมนาหนึ่งทศวรรษไอทีจุฬา : การเพิ่มศักยภาพระบบการจัดการด้วยเทคโนโลยีสารสนเทศ. จุฬาลงกรณ์มหาวิทยาลัย.
- [3] สำนักงานส่งเสริมสวัสดิภาพและพิทักษ์เด็ก เยาวชน ผู้ด้อยโอกาสและผู้สูงอายุ, รายงานเกี่ยวกับความรุนแรงต่อเด็ก, กระทรวงพัฒนาสังคมและความมั่นคงของมนุษย์, 2555.
- [4] R. Agrawal and R. Srikant. Fast Algorithms for Mining Association Rules. In Proceedings of the 20th VLDB Conference, Santiago, Chile, 1994.
- [5] Arnon Rungsawang, Athichart Tangpong, Pawat Laohawee and Tawa Khampachua, Novel Query Expansion Technique using Apriori Algorithm, The Eighth Text REtrieval Conference (TREC 8), Gaithersburg, USA, November 1999

การศึกษาเปรียบเทียบคุณภาพเสียงวีโอไอพีระหว่างแอปพลิเคชันไลน์ กับ เฟสบุ๊ก ด้วยการประเมินเชิงจิตวิสัย

A Comparative Study of VoIP Quality between LINE Application and Facebook Using Subjective Assessment

เขมิกา มุสิกพันธ์ (Khemika Moosikaphan)¹, เทอดพงษ์ แดงสี (Therdpong Daengsi)²

และ พงษ์พิสิฐ วุฒิชัยรุ โชติ (Pongpisit Wuttidittachotti)³

^{1,3} คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ กรุงเทพฯ

² เอ็นเตอร์ไพรส์เซอร์วิสเชส บริษัท แจ็คส์คอม จำกัด กรุงเทพฯ

mikaz@live.com, Therdpong1@yahoo.com, pongpisitw@kmutnb.ac.th

บทคัดย่อ

บทความนี้ได้ศึกษาเปรียบเทียบคุณภาพเสียง VoIP ระหว่าง แอปพลิเคชันไลน์ (LINE) กับ เฟสบุ๊ก (Facebook) โดยใช้การประเมินคุณภาพเสียงด้วยวิธีการสนทนาถึงสัมภพณ์ ซึ่งเป็นวิธีการเชิงจิตวิสัย แล้วนำค่าคุณภาพเสียงหรือค่าเอ็มโอเอส (Mean Opinion Score : MOS) ที่ได้มาวิเคราะห์เชิงสถิติด้วย ANOVA และ T-test ซึ่งพบว่า LINE ที่เชื่อมต่อผ่านเครือข่าย LAN ให้คุณภาพเสียงที่ดีที่สุด (MOS = 3.88) ในขณะที่ Facebook ที่เชื่อมต่อผ่านเครือข่าย WLAN ให้คุณภาพเสียงต่ำที่สุด (MOS = 2.75) และพบว่าคุณภาพเสียงที่ได้จาก LINE ดีกว่าคุณภาพเสียงที่ได้จาก Facebook อย่างมีนัยสำคัญ (P-Value < 0.05) ไม่ว่าเชื่อมต่อผ่านเครือข่ายแบบ LAN หรือ WLAN

คำสำคัญ: วีโอไอพี, จิตวิสัย, เอ็มโอเอส, ไลน์, เฟสบุ๊ก

Abstract

This paper presents a comparison of perceptual VoIP quality between LINE Application and Facebook using subjective assessment, call interview-conversation tests. The Mean Opinion Score (MOS) results were analyzed by using ANOVA and T-test. It has been found that LINE with Local Area Network (LAN) provides the best voice quality (MOS=3.88), whereas Facebook with Wireless LAN (WLAN) provides the worst voice quality (MOS=2.75). Also, it has been found that voice quality from LINE is better than voice quality from Facebook

with significant different (P-Value < 0.05), both LAN and WLAN connections.

Keyword: VoIP, Subjective, MOS, LINE, Facebook

1. บทนำ

เทคโนโลยี VoIP (Voice over Internet Protocol) พัฒนาขึ้นเพื่อการสื่อสารโดยวิธีส่งแพ็กเก็ตเสียงถึงกันผ่านเครือข่ายอินเทอร์เน็ต เป็นอีกทางเลือกหนึ่งที่สะดวกในการสื่อสารและช่วยลดค่าใช้จ่ายโทรศัพท์ [1] แต่การสื่อสารแบบนี้มีข้อจำกัดที่ส่งผลกระทบต่อคุณภาพเสียง เช่น เกิดการสูญเสียของแพ็กเก็ตเสียงที่ส่งถึงกัน ทำให้ข้อมูลสูญหาย (Loss) บางครั้งเกิดการประวิงเวลา (Delay) ทำให้เกิดผลกระทบต่อคุณภาพเสียงที่ได้ วิธีการวัดคุณภาพเสียงวีโอไอพีสามารถวัดได้ด้วยมาตรวัดเอ็มโอเอส หรือ MOS (Mean Opinion Score) ตามวิธีของไอทียู-ที (ITU-T) [2-3]

ซึ่งบทความนี้ได้ศึกษาเปรียบเทียบคุณภาพเสียงของบริการ VoIP 2 ชนิดที่ได้รับความนิยมสูงสุดในการ Download จากแอปพลิเคชันสโตร์ (Application Store) คือ ไลน์ (LINE) และ เฟสบุ๊ก (Facebook) โดยใช้วิธีการประเมินเชิงจิตวิสัย สำหรับการประมาณค่า MOS ใช้ข้อมูลจากกลุ่มผู้ใช้คนไทย ผ่านเครือข่าย LAN และ WLAN ของผู้ให้บริการอินเทอร์เน็ต True แล้วนำค่า MOS ที่ได้มาวิเคราะห์เชิงสถิติ เพื่อพิสูจน์สมมติฐานว่าคุณภาพเสียง VoIP ที่วัดได้มีความแตกต่างกันอย่างมีนัยสำคัญหรือไม่

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 VoIP

VoIP เป็นการนำเอาสัญญาณเสียงรวมเข้ากับสัญญาณข้อมูลเพื่อส่งผ่านไปยังระบบเครือข่ายผ่านทาง Internet Protocol แต่ก็ยังมีจุดด้อย คือ ในบางกรณีคุณภาพเสียงอาจไม่ดีเท่าโทรศัพท์และอาจมีการประวิงเวลา (Delay) หรือสัญญาณเสียงเดินทางมาถึงช้า ทำให้พูดสวนกันไม่ได้ไม่ถนัด และไม่สามารถใช้งานโทรศัพท์ VoIP ได้เมื่อไฟฟ้านดับ หรือ Internet เกิดขัดข้อง

2.2 วิธีการวัดคุณภาพเสียงสำหรับ VoIP

วิธีการวัดคุณภาพเสียงสำหรับ VoIP จะแบ่งออกเป็น 2 ประเภท [4] คือ การวัดเชิงวัตถุวิสัย เป็นการประเมินคุณภาพเสียงโดยใช้เครื่องมือในการวัด และการวัดเชิงจิตวิสัย คือการประเมินคุณภาพเสียงที่ใช้นั้นเป็นผู้ประเมิน โดยผลลัพธ์ของวิธีนี้สามารถนำมาเป็นเกณฑ์มาตรฐานสำหรับวิธีการวัดเชิงวัตถุวิสัย นอกจากนี้ยังมีนักวิจัยอีกหลายท่านระบุไว้ว่าการวัดคุณภาพเสียงเชิงจิตวิสัยเป็นวิธีการที่มีความแม่นยำสูงและน่าเชื่อถือดังหลักฐานที่ปรากฏต่อไปนี้

“การวัดคุณภาพเสียงเชิงจิตวิสัยถือว่าเป็นวิธีที่ “น่าเชื่อถือ” ที่สุดสำหรับวิธีการวัดคุณภาพเสียง” [5]

“การวัดคุณภาพเสียงเชิงจิตวิสัยเป็นวิธีที่เชื่อถือได้มากที่สุดสำหรับการวัดคุณภาพเสียงของผู้ฟังและมีผลดีในแง่ของความสัมพันธ์คุณภาพเสียงที่เป็นจริง” [6]

“...วิธีการวัดคุณภาพเสียงเชิงจิตวิสัยเป็นวิธีการที่ถูกต้องที่สุดสำหรับการวัดคุณภาพเสียงซึ่งจะใช้ในการเปรียบเทียบกับวิธีการวัดคุณภาพเสียงเชิงวัตถุวิสัย...” [7]

2.3 MOS (Mean Opinion Score)

MOS (Mean Opinion Score) เป็นวิธีวัดคุณภาพเสียงที่ถูกกำหนดเป็นมาตรฐานโดย ITU-T [8] ซึ่งวิธีนี้ทดสอบโดยใช้ผู้ทดสอบจำนวนมากให้โทรศัพท์พูดคุยซึ่งกันและกัน และบุคคลเหล่านั้นต้องมีทักษะในการฟังที่ดีพอสมควร จากนั้นให้คะแนนคุณภาพของเสียงที่ได้ยิน ผลการทดสอบจะมาจากผู้ที่เข้าร่วมทดสอบทุกคน โดย MOS เป็นค่าเฉลี่ยที่ได้จากความเห็นของผู้ใช้ที่กำหนดเป็นตัวเลข 1-5 ซึ่งแต่ละคะแนนจะมีความหมายดังตารางที่ 1 [8]

ตารางที่ 1 ค่าเอ็ม โอเอสและระดับคุณภาพเสียง

MOS	ระดับคุณภาพเสียง
5	ยอดเยี่ยม (Excellent)
4	ดี (Good)
3	พอใช้ (Fair)
2	ไม่ดี (Poor)
1	แย่มาก (Bad)

การวัดคุณภาพของเสียงสามารถวัดได้ในหลายลักษณะ [9] เช่น

2.3.1 คุณภาพเสียงที่ได้รับฟัง เป็นคุณภาพที่วัดจากเสียงที่ได้ฟังหลังจากผ่านเครือข่ายการสื่อสารแล้ว ผู้ฟังจะเป็นคนที่บอกได้ว่าคุณภาพเสียงที่ได้ยินเป็นอย่างไร ชัดเจน เหมือนเสียงต้นฉบับหรือไม่

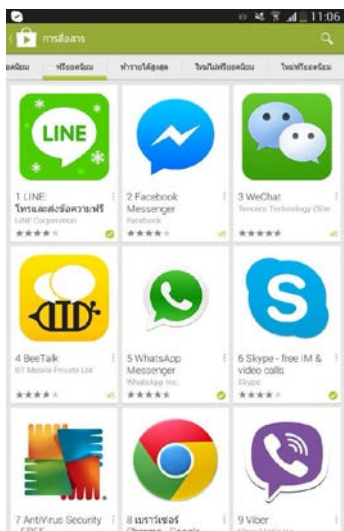
2.3.2 คุณภาพในการสนทนา รวมถึงคุณภาพเสียงที่ได้รับฟังและความสามารถในการตอบโต้สนทนาระหว่างการสื่อสาร ดังนั้น เรื่องของเสียงสะท้อน ดีเลย์ หรือปัญหาอื่นใดที่ทำให้รำคาญใจในการสนทนามีผลทั้งสิ้น

2.3.3 คุณภาพการเชื่อมต่อของเครือข่าย ไม่ว่าจะเป็นเรื่อง การเชื่อมต่อนั้นมีความเสถียรเพียงใด หรือมีการหลุดมามากน้อยเพียงใด เป็นต้น

2.4 ไลน์ (LINE)

LINE เป็น Application ที่สามารถใช้งานได้บนโทรศัพท์มือถือที่มีระบบปฏิบัติการ iOS, Android, Windows Phone และสามารถใช้งานได้บนคอมพิวเตอร์ PC หรือ Mac โดย LINE สามารถใช้โทรศัพท์หากันฟรีแบบ 1 ต่อ 1 สามารถส่งวิดีโอและข้อความเสียงหากันได้ฟรี

LINE เติบโตอย่างรวดเร็วเมื่อเทียบกับบริการ Social Network อื่นๆ แล้ว ถือว่า LINE ได้เปรียบมาก เพราะแค่ 257 วัน มีผู้ใช้ถึง 20 ล้านคน และหลังจากนั้นเพียง 6 เดือน เพิ่มขึ้นถึง 35 ล้านคน เห็นได้ว่า LINE มีผู้ใช้เพิ่มขึ้นมากถึง 600% [10] ซึ่งตอนนี้ยังคงมียอด Download สูงขึ้นเรื่อยๆ จนกลายเป็น Application ฟรีอันดับหนึ่งใน Play Store ของ Google ดังแสดงในภาพที่ 1 (ข้อมูลล่าสุด ณ วันที่ 6 กุมภาพันธ์ 2557)



ภาพที่ 1 อันดับยอดนิยมในแอปพลิเคชันสตรี (6/02/2557)

2.5 เฟสบุ๊ค (Facebook)

Facebook คือ เว็บไซต์เครือข่ายสังคม (Social Network) บน Internet ที่ทำให้ผู้ใช้สามารถสื่อสารและทำกิจกรรมร่วมกับผู้ใช้ Facebook คนอื่นๆ ไม่ว่าจะเป็นการตั้งประเด็นถามตอบในเรื่องที่สนใจ เขียนบทความ โพสต์รูปภาพคลิปวิดีโอ แชทสนทนา ได้ตอบ หรือใช้โทรศัพท์หากันฟรี เนื่องจากสามารถส่งวิดีโอ และข้อความเสียงหากันได้ และยังสามารถใช้ Facebook บน โทรศัพท์มือถือที่มีระบบ ปฏิบัติการ iOS, Android, Windows Phone ได้ด้วย Facebook เปิดใช้งานเมื่อ 4 กุมภาพันธ์ ค.ศ. 2004 ดำเนินงานและมีเจ้าของคือ บริษัทเฟสบุ๊ค (Facebook, Inc.) จากข้อมูล 1 มกราคม 2557 Facebook มีผู้ใช้ประจำมากกว่าหนึ่งพันสามร้อยล้านบัญชี ซึ่งมากเท่ากับจำนวนประชากรของประเทศจีนทั้งประเทศ [11]

2.6 ระบบเครือข่ายท้องถิ่น (LAN) และระบบเครือข่ายท้องถิ่นไร้สาย (WLAN)

LAN คือระบบเครือข่าย แบบเชื่อมต่อคอมพิวเตอร์เข้าด้วยกันในระยะจำกัด โดยใช้สายสัญญาณเป็นตัวกลางในการเชื่อมต่อ อัตราเร็วของเครือข่าย LAN อยู่ที่ระหว่าง 1-100 Mbps ทั้งนี้ความเร็วข้อมูลขึ้นอยู่กับ ตัวกลางสายส่งที่ใช้ เทคนิคการส่งสัญญาณ และข้อกำหนดของผู้ให้บริการเน็ตเวิร์ค [12]

WLAN คือ ระบบการสื่อสารข้อมูลที่มีรูปแบบในการสื่อสารแบบไม่ใช้สาย โดยใช้การส่งคลื่นความถี่วิทยุในย่านวิทยุ RF และ คลื่นอินฟราเรด ในการรับและส่งข้อมูลระหว่างคอมพิวเตอร์แต่ละเครื่อง ผ่านอากาศ, ทะลุกำแพง, เพดานหรือ

สิ่งก่อสร้างอื่นๆ นอกจากนั้นระบบ WLAN ยังมีคุณสมบัติครอบคลุมทุกอย่างเหมือนกับระบบ LAN แต่มีกพบปัญหาบางประการในระบบ WLAN คือ ระบบขาดเสถียรภาพ ไม่สามารถบริหารจัดการระบบจากส่วนกลางได้ และ ยังมีประเด็นเรื่องการรองรับการทำ Convergence Network [13]

2.7 ทบทวนวรรณกรรม

งานวิจัยเกี่ยวกับการวัดคุณภาพเสียงวีโอไอพีหลายงานที่เลือกเปรียบเทียบคุณภาพเสียงวีโอไอพีด้วยการใช้มาตรวัด MOS ดั้งบทความของ C. Wen-Hui และคณะ ได้สังเกตเห็นว่าการสื่อสารแบบ VoIP มีราคาต่ำกว่าการให้บริการเสียงแบบดั้งเดิม PSTN (Public Switched) และ VoIP ยังมีความยืดหยุ่นที่มากขึ้นแต่คุณภาพของเสียงที่ดีนั้นก็ขึ้นอยู่กับ การเชื่อมต่อของอินเทอร์เน็ตที่ดีด้วย จึงได้ทำการทดลองประเมินผล แอปพลิเคชันวีโอไอพียอดนิยมได้แก่ MSN และ Skype โดยทดสอบด้วยค่า MOS ซึ่งสามารถนำผลที่ได้มาปรับปรุงการส่งผ่านข้อมูลของ Skype ได้ถึง 50% และยังมี ความเชื่ออีกว่าการศึกษานี้จะสามารถนำไปออกแบบบริการเสียงที่ดีขึ้นต่อไป [14]

S. Batu และ W. Benjamin ได้ประเมินคุณภาพเสียง VoIP สำหรับ VoIP Application ยอดนิยมในช่วงเวลานั้น ได้แก่ Skype, Google Talk, Windows Live Messenger และ Yahoo Messenger ด้วยเครื่องวัด PESQ ซึ่งเป็นวิธีวัดเชิงจิตวิสัยภายใต้เงื่อนไขความล่าช้าของเครือข่าย ซึ่งผลปรากฏว่า Windows Live Messenger มีประสิทธิภาพมากกว่า Application อื่นๆ เมื่อมีการสูญเสียแพ็คเกจเกิดขึ้น โดยผลของการวิจัยครั้งนี้เป็นที่น่าสนใจว่า Windows Live Messenger เป็นที่นิยมอันดับหนึ่งเมื่อเทียบกับ Application อีก 3 ชนิดที่ทำกรวิจัย [15] นอกจากนี้ S. Batu และ W. Benjamin ยังได้ประเมินคุณภาพเสียงของระบบ VoIP ระหว่าง Skype และ Google Talk โดยใช้ UDP Packet ใช้การกำหนดวิธีการเข้ารหัสการพูด ควบคุม Jitter และค่าการสูญเสียแพ็คเกจ ทำให้ทราบว่า Skype มีการปรับตัวภายใต้สภาวะต่างๆที่กำหนด แต่ Google Talk ไม่ได้มีการปรับตัวตามสภาพการใช้งานหรือผลกระทบต่างๆ ที่เกิดขึ้น [16]

B. Rodrigo และคณะ ได้ประเมินคุณภาพการสื่อสารด้วยเสียงของ Skype และ Google Talk ภายใต้สภาพเครือข่ายที่มีการสูญเสียแพ็คเกจเพื่อประเมินคุณภาพของเสียงระบบ VoIP

ชี้ให้เห็นจุดอ่อนจุดแข็งของระบบ และจัดอันดับสำหรับแอปพลิเคชัน VoIP ที่สื่อสารแบบ P2P โดยใช้เครื่องวัด PESQ และเปรียบเทียบอัลกอริทึมเดิมกับ ITU ซึ่งผลการทดสอบทำให้ทราบว่า Skype มีคุณภาพเสียงที่ดีกว่าในสภาวะที่เกิดการสูญเสียแพ็คเกจ และยังมีผลคลาดเคลื่อนเกิดขึ้นทำให้สรุปได้ว่า PESQ-MOS ไม่ได้เป็นอัลกอริทึมที่เหมาะสมที่สุดในการประเมินคุณภาพเสียง [17]

B. Dario และคณะได้วิเคราะห์รายละเอียดการจราจรของข้อมูลสำหรับ Skype ซึ่งเป็น VoIP Application ที่เป็นที่ยอมรับในขณะนั้น ศึกษาทั้งในลักษณะของการสื่อสารด้วยเสียง และวิดีโอ ภายใต้เงื่อนไขการสื่อสารบน Protocol แบบ TCP และ UDP วัดค่า Testbed ทำให้ทราบว่า การตอบสนองของ Skype บน Protocol ทั้ง 2 แบบ แตกต่างกันและการจราจรของข้อมูลสำหรับ Skype นั้นมีการเกิด Floods เพียงระยะสั้นๆ เมื่อมีการส่งข้อความให้กับผู้ใช้หลายคนในระบบซึ่งข้อมูลต่างๆ ที่ค้นพบสามารถนำไปพัฒนาปรับปรุงระบบ Skype ให้ดีขึ้นต่อไป [18]

C. Wu และคณะได้ประเมินขนาด Buffer ของ Skype, Google Talk และ MSN Messenger ซึ่งพบว่า MSN Messenger มีการปรับขนาดของ Buffer ได้เหมาะสมกับสภาพการเชื่อมต่อของเครือข่ายดีที่สุด ในขณะที่ Skype ไม่มีวิธีการปรับขนาด Buffer ในการเชื่อมต่อที่ต่างกันเลย [19]

P. Casa ได้ศึกษาเปรียบเทียบเว็บไซต์เครือข่ายสังคมที่เป็นที่นิยมอยู่ในมือถือได้แก่ YouTube และ Facebook พบว่า YouTube QoE ตอบสนองต่อ Downlink-encoding ได้ ในขณะที่ Facebook ตอบสนองได้ช้า แต่น่าเสียดายที่การศึกษาครั้งนี้ไม่ได้รายงานเกี่ยวกับปัญหาคุณภาพเสียง [20]

N. Chumkot และคณะ ได้เปรียบเทียบคุณภาพเสียง VoIP ด้วยวิธีการวัดแบบ PESQ ร่วมกับการวิเคราะห์เชิงสถิติด้วย ANOVA และ T-test ผ่านเครือข่ายไอพีและ 3G สำหรับ LINE และ Tango โดยใช้ภาษาจีนภาษาไทย และภาษาอังกฤษ ทำให้ทราบว่า LINE ให้คุณภาพเสียงดีกว่า Tango ทั้ง 3 ภาษา และคุณภาพเสียงภายใต้เครือข่าย 3G ดีกว่า คุณภาพเสียงภายใต้เครือข่ายไอพีอย่างมีนัยสำคัญ [21]

จากงานวิจัยที่ได้กล่าวมาข้างต้น จะเห็นได้ว่า ยังไม่มีผู้วิจัยที่ทำการศึกษาระบบเปรียบเทียบคุณภาพเสียง VoIP ของ LINE และ Facebook ด้วยวิธีการประเมินเชิงจิตวิสัย บนเครือข่าย LAN และ WLAN

3. การทดสอบ

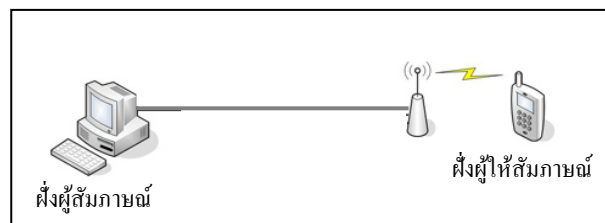
3.1 การออกแบบการทดสอบและการเก็บข้อมูล

ติดตั้ง LINE และ Facebook บนเครื่องคอมพิวเตอร์ที่เตรียมไว้แล้วติดต่อไปยังผู้ให้สัมภาษณ์เพื่อเก็บข้อมูล ดังภาพที่ 2

ควบคุมการทดสอบ โดยทดสอบผ่านเครือข่าย Internet ที่ให้บริการโดยผู้ให้บริการรายเดียวกัน ใช้อุปกรณ์ต่างๆ ชุดเดียวกัน ภายในห้องปฏิบัติการชั่วคราวห้องเดียวกัน ซึ่งวัดความเร็ว Internet ที่ใช้ในการทดสอบคุณภาพเสียง VoIP ผ่านเว็บไซต์ www.speedtest.net จำนวน 40 ครั้ง ทดสอบพร้อมๆ กับระยะเวลาที่ใช้เก็บข้อมูลช่วงเดือนธันวาคม 2556 ถึงเดือนมกราคม 2557 ได้ความเร็วเฉลี่ยของ Internet แสดงดังตารางที่ 2 และเก็บรวบรวมข้อมูลภายใต้เงื่อนไขที่กำหนดขึ้นดังตารางที่ 3

3.2 การประเมินคุณภาพเสียงเชิงจิตวิสัย

ใช้วิธีการทดสอบแบบสนทนากึ่งสัมภาษณ์ โดยผู้สัมภาษณ์จะสนทนาโต้ตอบกับผู้ให้สัมภาษณ์ และผู้สัมภาษณ์เป็นผู้บันทึกผลการประเมินลงในแบบประเมินผลการทดสอบคุณภาพเสียง VoIP ที่สร้างขึ้น แล้วนำค่า MOS ที่ได้มาวิเคราะห์เชิงสถิติด้วย ANOVA และ T-test



ภาพที่ 2 วิธีการเก็บข้อมูล

ตารางที่ 2 ความเร็วเฉลี่ยของ Internet ที่ใช้วัดได้ในช่วงเวลาใกล้เคียงกับการทดสอบ

	LAN		WLAN	
	Upload Speed/Mbps	Download Speed/Mbps	Upload Speed/Mbps	Download Speed/Mbps
Average	15.14	18.32	1.36	17.41
SD	0.79	0.62	0.15	1.47

ตารางที่ 3 เงื่อนไขการทดสอบเชิงจิตวิสัยสำหรับเก็บข้อมูล

	LAN	WLAN
LINE	24 คน (ชาย 10 คน, หญิง 14 คน)	24 คน (ชาย 13 คน, หญิง 11 คน)
Facebook	24 คน (ชาย 7 คน, หญิง 17 คน)	24 คน (ชาย 11 คน, หญิง 13 คน)

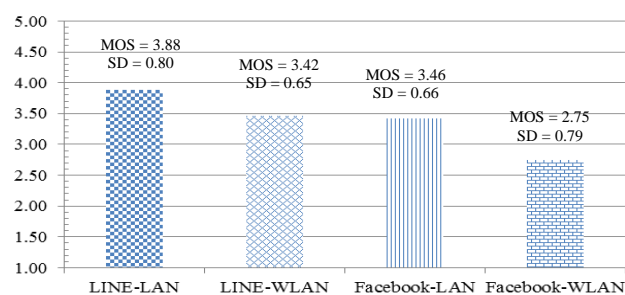
4. ผลการทดสอบ

หลังจากเก็บรวบรวมข้อมูลจากผู้ร่วมทดสอบจำนวน 96 คน แบ่งเป็นเพศชาย 41 คน (42.71%) เพศหญิง 55 คน (57.29%) ซึ่งมีอายุเฉลี่ยอยู่ที่ 22.68 ปี ค่า SD อายุเท่ากับ 4.68 ค่ารวมค่าเฉลี่ยของค่า MOS ได้ดังภาพที่ 3 ซึ่งจะเห็นได้ว่าผลลัพธ์ของคุณภาพเสียงของ LINE ดีกว่า Facebook ซึ่ง LINE ที่เชื่อมต่อผ่านเครือข่าย LAN ให้คุณภาพเสียงดีที่สุด (MOS = 3.88) และ Facebook ที่เชื่อมต่อผ่านเครือข่าย WLAN ให้คุณภาพเสียงต่ำที่สุด (MOS = 2.75)

5. การวิเคราะห์และอภิปราย

การทดสอบนี้เป็นการทดสอบเชิงจิตวิสัย ซึ่งเป็นวิธีการทดสอบที่ทดสอบกับผู้ใช้งานจริงไม่ใช่การทดสอบเชิงวัตถุวิสัย ซึ่งเป็นการวัดด้วยเครื่องมือ เช่น การวัดด้วย PESQ และ E-model [9] หลังการทดสอบ ได้วิเคราะห์ข้อมูลเชิงสถิติด้วย ANOVA และ T-test เพื่อพิสูจน์สมมติฐานว่าคุณภาพเสียงระหว่าง LINE และ Facebook ภายใต้การเชื่อมต่อด้วยเครือข่ายที่ต่างกัน มีคุณภาพเสียง VoIP แตกต่างกันอย่างมีนัยสำคัญหรือไม่ ซึ่งทำการทดสอบนัยสำคัญทางสถิติที่ 0.05 หรือที่ช่วงความเชื่อมั่น 95% โดยตั้งสมมติฐานเกี่ยวกับค่า MOS ที่วัดได้จาก Application ทั้ง 2 เป็นสมมติฐาน A,B,C,D และ E ดังตารางที่ 4

MOS-CQS



ภาพที่ 3 กราฟแสดงค่า MOS ที่ได้จากทดสอบ

ตารางที่ 4 แสดงค่า P-Value ที่ได้จากการพิสูจน์สมมติฐาน

สมมติฐาน	P-Value
A	<0.001*
H_0 : MOS from all conditions are the same H_1 : MOS from all conditions are different	
B	0.055
H_0 : MOS-LINE- LAN = MOS-LINE-WLAN H_1 : MOS-LINE- LAN \neq MOS-LINE-WLAN	
C	0.003*
H_0 : MOS-FB-LAN = MOS-FB-WLAN H_1 : MOS-FB-LAN \neq MOS-FB-WLAN	
D	0.035*
H_0 : MOS-LINE- LAN = MOS-FB-LAN H_1 : MOS-LINE- LAN \neq MOS-FB-LAN	
E	0.002*
H_0 : MOS-LINE-WLAN = MOS-FB-WLAN H_1 : MOS-LINE-WLAN \neq MOS-FB-WLAN	

หมายเหตุ * คือ มีนัยสำคัญทางสถิติที่ระดับ 0.05

จากตารางที่ 4 จะเห็นว่าสมมติฐาน A มีค่า P-Value < 0.001 แสดงให้เห็นว่าคุณภาพเสียงของ LINE กับ Facebook ที่เชื่อมต่อผ่านเครือข่ายแบบ LAN และ WLAN อย่างน้อย 1 เงื่อนไข ที่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ เมื่อเทียบกับผลที่ได้จากเงื่อนไขอื่น ๆ

สมมติฐาน B มีค่า P-Value = 0.055 แสดงให้เห็นว่าคุณภาพเสียงของแอปพลิเคชัน LINE ภายใต้การเชื่อมต่อแบบ LAN และ WLAN ไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

สมมติฐาน C มีค่า P-Value = 0.003 แสดงให้เห็นว่าคุณภาพเสียงของ Facebook ภายใต้การเชื่อมต่อแบบ LAN และ WLAN มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

สมมติฐาน D มีค่า P-Value = 0.035 และ สมมติฐาน E มีค่า P-Value = 0.002 แสดงให้เห็นว่าคุณภาพเสียงของ LINE กับ Facebook ภายใต้การเชื่อมต่อทั้งแบบ LAN และ WLAN แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

อย่างไรก็ตาม ค่า P-Value จากสมมติฐาน B ที่ทดสอบได้ มีค่า P-Value ใกล้เคียง 0.05 มาก นั้นหมายความว่า อาจเกิดความผิดพลาดแบบที่ 2 (Error Type II) คือการยอมรับสมมติฐานที่เป็นเท็จขึ้นได้ และ การทดสอบครั้งนี้เป็นเพียงการทดสอบเบื้องต้นจึงควรเก็บข้อมูลเพิ่มเติมหรือทดสอบซ้ำอีกครั้ง

6. บทสรุปและข้อเสนอแนะ

จากการศึกษาเปรียบเทียบคุณภาพเสียง VoIP ระหว่าง LINE และ Facebook ด้วยการประเมินจากผู้ร่วมทดสอบจำนวน 96 คน พบว่าคุณภาพเสียงที่ได้จากแอปพลิเคชัน LINE ดีกว่าคุณภาพเสียงที่ได้จาก Facebook อย่างมีนัยสำคัญ (P-Value < 0.05) ไม่ว่าจะเชื่อมต่อผ่านเครือข่ายแบบ LAN หรือ WLAN นอกจากนี้ยังพบว่า ทั้ง LINE และ Facebook ที่เชื่อมต่อผ่านเครือข่ายแบบ LAN ต่างก็ให้คุณภาพเสียงดีกว่าการเชื่อมต่อผ่านเครือข่ายแบบ WLAN

อย่างไรก็ตามเครือข่าย WLAN ก็ยังมีข้อดีคือมีความสะดวกต่อการใช้งาน และคุณภาพเสียงที่วัดได้ก็อยู่ในระดับที่ยอมรับได้ ซึ่งผลที่ได้จากการศึกษาครั้งนี้ สามารถนำไปใช้เป็นข้อเสนอแนะให้ผู้ใช้เลือกใช้เครือข่ายและบริการ VoIP ที่เหมาะสมเพื่อให้ได้คุณภาพเสียงที่ดีที่สุด

7. กิตติกรรมประกาศ

บทความฉบับนี้สำเร็จลุล่วงไปด้วยดี ด้วยความช่วยเหลือของเพื่อนๆในกลุ่ม LINE และ Facebook ตลอดจนเพื่อนๆในกลุ่มวิจัย VoIP คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

8. เอกสารอ้างอิง

- [1] ปรีชาภัทร มะณี, “การสื่อสารด้วยระบบ Voice Over IP” [ระบบออนไลน์] www.et.prm.chula.ac.th, 2013.
- [2] บรรจบ สุขประภากรณ, “เทคโนโลยีการสื่อสารโทรศัพท์ด้วยระบบ VoIP ” [ระบบออนไลน์] <http://www.ind.cru.ac.th> , 2013.
- [3] ITU-R, “Recommendation ITU-R M.1645, Framework and overall objectives of the future development of IMT-2000 and system beyond IMT-2000” 2003.
- [4] ITU-T Recommendation P.805, “Subjective evaluation of conversational quality.” Apr. 2007.
- [5] M. Narbutt and M. Davis, “Assessing the quality of VoIP transmission affected by playout buffer scheme,” in proc. Of 4th MESAQIN 2005, Prague, Czech Republic. Jun. 2005
- [6] M. Goudarzi, “Evaluation of Voice Quality in 3G Mobile Networks,” Thesis, University of Plymouth, Plymouth, UK, Jun. 2008.
- [7] M. Al-Akhras, H. Zedan, R. John and I. Almomani, “Non-intrusive speech quality prediction in VoIP networks using a neural network approach,” Neurocomputing, vol. 72, pp. 2595-2608, 2009.
- [8] ITU-T Recommendation P.800, “Methods for Subjective Determination of Transmission Quality”, 1996
- [9] F. De Rango et al., “Overview on VoIP: Subjective and Objective Measurement Methods” IJCSNS International Journal of Computer Science and Network Security, Vol.6,No.1B, Jan. 2006.
- [10] LINE, “LINE Exceeds 300 Million Users Worldwide,” Available from : <http://linecorp.com/en/press/2013/1125635> [Accessed Jan. 2014].
- [11] Statistic Brain, “Facebook Statistics,” Available from : <http://www.statisticbrain.com/facebook-statistics> [Accessed Jan. 2014].
- [12] Webopedia, “Local-Area Network” Available from : http://www.webopedia.com/TERM/L/local_area_network_LAN.html [Accessed Jun. 2013].
- [13] A. Ksentini, “Building the bridges between QoS and QoE for network control mechanisms,” Available from : http://hal.archives-ouvertes.fr/docs/00/91/38/72/PDF/these_HDR_ksentini.pdf [Accessed Jan. 2014]
- [14] C. Wen-Hui, X. Wei-Cheng and C. Cheng-Fu., “A Performance Study of VoIP Application : MSN vs. Skype” The MULTICOMM 2006.
- [15] S. Batu and W. Benjamin., “Evaluation Of Conversational Voice Communication Quality Of The Skype, Google-Talk, Windows Live and Yahoo Messenger VoIP Systems” IEEE Workshop On Multimedia Signal Processing, 2007.
- [16] S. Batu and W. Benjamin., “Analysis and Evaluation of The Skype and Google-Talk VoIP Systems” IEEE INT’L Conference On Multimedia and Expo, 2006.
- [17] B. Rodrigo., et al “Performance evaluation of P2P VoIP applications” The 17th International Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV 2007), Jun. 2007.
- [18] B. Dario., et al “Detailed Analysis of Skype Traffic” IEEE Transactions On Multimedias, Vol.11, No.1, Jan. 2009, p.117-127
- [19] C.-C. Wu, K.T. Chen, C.-Y. Huang, and C.-L. Lei, “An Empirical Evaluation of VoIP Playout Buffer Dimensioning in Skype, Google Talk, and MSN Messenger,” Proc. NOSSDAV, Williamsburg, VA, Jun. 2009, P. 97-102.
- [20] P. Casa, A. Sackl, S. Egger, and R. Schatz, “YouTube & Facebook Quality of Experience in Mobile Broadband Networks,” Proc. IEEE Globecom Workshops, Anaheim, CA, Dec. 2012, p.1269-1274.
- [21] นฤมล ชุมคช และคณะ “การศึกษาเปรียบเทียบคุณภาพเสียงวีโอไอพีด้วยวิธีการวัดแบบพีอีเอสคิว : กรณีศึกษาแอปพลิเคชันไลน์ กับแท่งโก้ ผ่านเครือข่ายไอพีของ มจพ. และเครือข่าย 3G” การประชุมวิชาการทางคอมพิวเตอร์และเทคโนโลยีสารสนเทศ CIT&UniNoMS สระแก้ว ประเทศไทย, 31 มกราคม 2557 หน้า 70-75.

การจัดตารางงานบนพื้นฐานของการกระจายงานสมดุลในการประมวลผล แบบกลุ่มเมฆโดยขั้นตอนวิธีอาณานิคมผึ้งเทียม Task Scheduling Based on Load Balancing in Cloud Computing Using Artificial Bee Colony Algorithm

บุญหทัย เครือแก้ว (Boonhatai Kruekaew)¹ และวารงคณา กิมปาน (Warangkha Kimpan)²

สาขาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

¹s3650853@kmitl.ac.th, ²knwarang@kmitl.ac.th

บทคัดย่อ

เครื่องคอมพิวเตอร์เสมือนถือเป็นส่วนประกอบสำคัญของเทคโนโลยีการประมวลผลแบบกลุ่มเมฆ ดังนั้นการจัดตารางงานเพื่อให้เครื่องคอมพิวเตอร์เสมือนมีการกระจายงานแบบสมดุลและระบบการประมวลผลแบบกลุ่มเมฆสามารถทำงานได้อย่างมีประสิทธิภาพจึงเป็นปัญหาที่สำคัญ บทความนี้ได้เสนอการใช้ขั้นตอนวิธีอาณานิคมผึ้งเทียมควบคู่กับการจัดลำดับงานแบบฮิวริสติก สำหรับจัดตารางงานในการเข้าใช้เครื่องคอมพิวเตอร์เสมือน ทั้งในกรณีที่เครื่องคอมพิวเตอร์เสมือนมีสภาพแวดล้อมภายในเหมือนกันและต่างกันบนการประมวลผลแบบกลุ่มเมฆ โดยมีจุดมุ่งหมายเพื่อหาขั้นตอนวิธีในการกระจายงานแบบสมดุลให้กับเครื่องคอมพิวเตอร์เสมือนที่มีสภาพแวดล้อมต่างกันและใช้เวลารวมในการทำงานน้อยที่สุด การทดลองได้ทำการเปรียบเทียบเพื่อหาวิธีการที่เหมาะสมสำหรับการจัดตารางงานในการเข้าใช้เครื่องคอมพิวเตอร์เสมือนโดยใช้ CloudSim ในการสร้างแบบจำลอง

คำสำคัญ: การประมวลผลแบบกลุ่มเมฆ การกระจายงานแบบสมดุล การจัดตารางงาน ขั้นตอนวิธีอาณานิคมผึ้งเทียม วิธีการจัดลำดับงานแบบฮิวริสติก

Abstract

Virtual Machine is a key component of Cloud computing Technology. Therefore, the developing an optimal scheduling mechanism for balancing Virtual Machines operations at Cloud computing framework is an intriguing issue for Cloud computing services performance. This paper proposed Artificial Bee Colony

(ABC) algorithm and Heuristic scheduling algorithms which are applied to optimize the scheduling tasks for Virtual Machines on homogeneous and heterogeneous environment platforms on Cloud computing. The main contribution of this work is to analyze the difference of Virtual Machines load balancing algorithms. The experimental results demonstrated the comparison of many algorithms to find the suitable one for task scheduling on Virtual Machines and to reduce the makespan of data processing time. The scheduling strategy was simulated by using CloudSim tool.

Keyword: Cloud Computing, Load Balance, Scheduling, Artificial Bee Colony Algorithm, Heuristics.

1. บทนำ

การประมวลผลแบบกลุ่มเมฆ (Cloud Computing) [1] ถือเป็นระบบคอมพิวเตอร์ที่มีการประมวลผลในระบบอินเทอร์เน็ต มีรูปแบบการประมวลผลที่อิงตามความต้องการของผู้ใช้ ผู้ใช้สามารถระบุความต้องการไปยังซอฟต์แวร์ของระบบ จากนั้นซอฟต์แวร์จะมีการร้องขอให้ระบบจัดสรรทรัพยากรและบริการให้ตรงกับความต้องการของผู้ใช้ นอกจากนี้ผู้ใช้งานสามารถเพิ่มและลดจำนวนทรัพยากรได้ตามความต้องการ โดยไม่จำเป็นต้องทราบถึงรูปแบบการทำงานหรือการให้บริการ ข้อดีของการประมวลผลแบบกลุ่มเมฆ คือ การลดความซับซ้อนของผู้ให้บริการ ประหยัดพลังงานและลดค่าใช้จ่ายในการจัดการ เพราะระบบจะทำงานผ่านเทคโนโลยีเสมือน (Virtualization Technology) ส่งผลให้ระบบจะไม่ถูกจำกัดในเรื่องของขีดความสามารถในการประมวลผล

เทคโนโลยีเสมือนเป็นการจำลองการสร้างเครื่องคอมพิวเตอร์เสมือน (Virtualization Machine: VM) [2,3] ด้วยซอฟต์แวร์ที่ทำให้คอมพิวเตอร์เครื่องหนึ่งสามารถทำงานเสมือนมีหลายระบบ โดยแต่ละเครื่องคอมพิวเตอร์เสมือนจะมีทรัพยากรที่เป็นอิสระต่อกัน ซึ่งถือเป็นการลดปริมาณเครื่องคอมพิวเตอร์ที่ใช้เป็นเครื่องแม่ข่าย (Server) และเป็นการใช้ทรัพยากรอย่างคุ้มค่า

ในกรณีที่มีการร้องขอเข้าใช้บริการในระบบไปยังเครื่องแม่ข่ายเดียวกัน ในบางครั้งเครื่องแม่ข่ายอื่นอาจไม่มีการเข้าใช้บริการ ซึ่งเรียกได้ว่าระบบเกิดการกระจายงานแบบไม่สมดุล (Load Imbalance) สามารถแก้ไขได้โดยการจัดตารางงาน (Scheduling) ก่อนเข้าใช้บริการ ซึ่งการจัดตารางงานที่ได้นั้นจะมีประโยชน์ในแง่ของการเพิ่มประสิทธิภาพการใช้ทรัพยากรคือ ลดเวลารอเข้าทำงานในระบบและระบบจะมีการกระจายการทำงานแบบสมดุล (Load Balance)

การจัดตารางงานในระบบประมวลผลแบบกลุ่มเมฆ เพื่อให้ระบบเกิดการกระจายงานแบบสมดุลนั้น จำเป็นต้องคำนึงถึงเงื่อนไขต่างๆ เช่น ความเสถียรของระบบ ความรวดเร็วในการทำงาน การเปลี่ยนแปลงของสภาพแวดล้อมในระบบ เป็นต้น เงื่อนไขที่ต้องคำนึงถึงในการจัดตารางงานนั้นจะคล้ายคลึงกับการทำงานของขั้นตอนวิธีอาณานิคมผึ้งเทียม (Artificial Bee Colony Algorithm: ABC Algorithm) ส่วนของพฤติกรรม การหาอาหารของฝูงผึ้ง เนื่องจากผึ้งสามารถปรับตัวได้ตามสภาพแวดล้อมที่อาศัยอยู่ [4,5] การประมวลผลแบบกลุ่มเมฆมีความคล้ายคลึงกับขั้นตอนวิธีอาณานิคมผึ้งเทียม ในส่วนของเป้าหมายของการทำงาน เนื่องจากระบบการประมวลผลแบบกลุ่มเมฆต้องการให้เครื่องคอมพิวเตอร์สามารถประมวลผลงานให้ได้ปริมาณมากที่สุด ในขณะที่ผึ้งต้องการแหล่งอาหารที่มีปริมาณน้ำหวานมากที่สุดเช่นกัน

ดังนั้นงานวิจัยนี้ได้นำเสนอขั้นตอนวิธีการจัดตารางงานเพื่อเข้าใช้เครื่องคอมพิวเตอร์เสมือน ในกรณีที่เครื่องคอมพิวเตอร์เสมือนมีสภาพแวดล้อมภายในเหมือนกันและต่างกัน (Homogeneous and Heterogeneous) โดยประยุกต์ใช้ขั้นตอนวิธีอาณานิคมผึ้งเทียมควบคู่กับการจัดลำดับการทำงานแบบฮิวริสติก (Heuristics) [6] เพื่อหาเวลารวมน้อยที่สุดในการทำงาน (Processing Time) และทำให้ระบบอยู่ในสภาวะสมดุล

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การประมวลผลแบบกลุ่มเมฆ

การประมวลผลแบบกลุ่มเมฆ มีรูปแบบการให้บริการที่นำเสนอทรัพยากรสารสนเทศ สำหรับการประมวลผลตั้งแต่โครงสร้างพื้นฐานไปจนถึงชุดคำสั่งประยุกต์ เสมือนหนึ่งเป็นบริการที่ผู้ใช้สามารถบริการตนเองได้ โดยเลือกใช้เฉพาะสิ่งที่ตนต้องการ ผู้ใช้บริการระบบการประมวลผลแบบกลุ่มเมฆจะมองกลุ่มเมฆแตกต่างกันออกไป เนื่องจากบริการที่นำเสนอในแต่ละผู้ให้บริการนั้นมีความแตกต่างกันออกไป รวมทั้งผู้ให้บริการสามารถเลือกใช้บริการเฉพาะที่ตนต้องการได้ [1]

การให้บริการบนระบบการประมวลผลแบบกลุ่มเมฆสามารถแบ่งรูปแบบของการให้บริการเป็น 3 รูปแบบ [4] คือ

1) การบริการโครงสร้างพื้นฐานของคอมพิวเตอร์ (Infrastructure as a Service: IaaS) มีประโยชน์ในการประมวลผลทรัพยากรจำนวนมาก ผู้ให้บริการโครงสร้างพื้นฐาน เช่น Amazon's EC2 เป็นต้น

2) การบริการแพลตฟอร์ม (Platform as a Service: PaaS) เป็นการประมวลผลซึ่งจะมีระบบปฏิบัติการและสนับสนุนเว็บแอปพลิเคชันเข้าร่วมด้วย เช่น Google Apps เป็นต้น

3) การบริการซอฟต์แวร์ (Software as a Service: SaaS) จะให้บริการการประมวลผลผ่านแอปพลิเคชันแม่ข่ายของผู้ให้บริการและเปิดให้บริการทางด้านซอฟต์แวร์ต่างๆ โดยผู้ใช้สามารถเรียกใช้งานซอฟต์แวร์ผ่านระบบอินเทอร์เน็ต

ในระบบการประมวลผลแบบกลุ่มเมฆนั้นจะมีการจัดสรรทรัพยากรของระบบแตกต่างกัน ขึ้นอยู่กับความต้องการของระบบ กรณีที่มีการจัดสรรทรัพยากรเหมือนกันทั้งระบบ จะเรียกสภาพแวดล้อมภายในระบบนี้ว่า Homogeneous และในกรณีที่มีการจัดสรรทรัพยากรในระบบแตกต่างกัน จะเรียกสภาพแวดล้อมนี้ว่า Heterogeneous

ตลอดระยะเวลาหลายปีที่ผ่านมาได้มีผู้เสนอแนะขั้นตอนวิธีการจัดตารางงานเข้าใช้ทรัพยากรในระบบการประมวลผลแบบกลุ่มเมฆด้วยวิธีการต่างๆ เพื่อให้ระบบมีการกระจายการทำงานแบบสมดุลและตรงกับความต้องการของผู้ใช้ ภายใต้ทรัพยากรที่มีอยู่อย่างจำกัด และเพื่อเป็นการเพิ่มประสิทธิภาพการทำงานของระบบการประมวลผลแบบกลุ่มเมฆ [3] เช่น

ในปี 2010 Y. Fang, F. Wang และ J. Ge [7] ได้เสนอการจัดตารางการทำงานในการประมวลผลแบบกลุ่มเมฆ โดยคำนึงถึงความต้องการของผู้ใช้งานที่มีการเปลี่ยนแปลงตลอดเวลาและคำนึงถึงการกระจายงานแบบสมดุลในสภาพแวดล้อมของการประมวลผลแบบกลุ่มเมฆ เมื่อผู้ใช้งานและความต้องการในการใช้งานทรัพยากรมีปริมาณเพิ่มมากขึ้น โดยใช้ CloudSim [8] ซึ่งเป็นชุดเครื่องมือที่ใช้ในการสร้างแบบจำลองของระบบการประมวลผลแบบกลุ่มเมฆ สามารถใช้ทดสอบประสิทธิภาพของระบบที่พัฒนาขึ้นใหม่ เป็นเครื่องมือที่สะดวกในการกำหนดสภาพแวดล้อมตามที่ผู้ใช้งานต้องการ เช่น การตั้งค่าของดาต้าเซ็นเตอร์ (Data Center) และเครื่องคอมพิวเตอร์เสมือน เป็นต้น

จากนั้นในปี 2011 B. Mondal และ K. Dasgupta [9] ได้มีการนำขั้นตอนการแก้ไขปัญหาการปีนเขาของสโตแคสติก (Stochastic Hill Climbing) มาใช้ในการจัดการงานเพื่อให้มีการกระจายงานแบบสมดุลในการประมวลผลแบบกลุ่มเมฆ

2.2 การจัดลำดับงานแบบฮิวริสติก

วิธีนี้เป็นที่น่าทึ่งต่างๆ มาใช้ในการหาผลลัพธ์ที่น่าพอใจของปัญหา เช่น กฎการจัดลำดับความสำคัญ วิธีการสุ่ม เป็นต้น วิธีการนี้อาจได้ผลลัพธ์ที่น่าพอใจ แต่ไม่สามารถรับรองได้ว่าเป็นผลลัพธ์ที่ดีที่สุด ในบทความนี้ได้นำกฎการจัดลำดับงานโดยใช้ลำดับความสำคัญเป็นเกณฑ์ในการเลือกขั้นตอนการทำงาน [6] ได้แก่

- 1) การจัดให้งานมาก่อนทำก่อน (First Come First Serve: FCFS) เป็นการจัดลำดับงานโดยทำงานที่เข้ามาก่อนเป็นอันดับแรกและทำงานที่เข้ามาหลังเป็นลำดับถัดไป
- 2) การจัดให้งานที่มีขนาดเล็กที่สุดทำงานก่อน (Shortest Job First: SJF) เป็นการจัดลำดับงานโดยทำงานที่มีขนาดเล็กที่สุดก่อน และเลือกงานที่มีขนาดใหญ่เป็นลำดับถัดไป
- 3) การจัดให้งานที่มีขนาดใหญ่ที่สุดทำก่อน (Longest Job First: LJF) เป็นการจัดลำดับงานโดยทำงานที่มีขนาดใหญ่ที่สุดก่อน และเลือกงานที่มีขนาดเล็กเป็นลำดับถัดไป

2.3 ขั้นตอนวิธีอาณานิคมผึ้งเทียม

ขั้นตอนวิธีอาณานิคมผึ้งเทียม [4,5] เป็นขบวนการในการหาค่าที่เหมาะสม โดยเลียนแบบจากพฤติกรรมในการหาอาหารของฝูงผึ้ง และเป็นกระบวนการหาคำตอบประเภทที่อยู่ในกลุ่ม

ที่มีแนวคิดมาจากการเลียนแบบพฤติกรรมของฝูงแมลง ได้มีการนำขั้นตอนวิธีอาณานิคมผึ้งเทียมไปประยุกต์ใช้ในการแก้ไขปัญหาต่างๆ เช่น การค้นหาค่าที่เหมาะสม การทำเหมืองข้อมูล เป็นต้น

ขั้นตอนวิธีอาณานิคมผึ้งเทียมเป็นขั้นตอนวิธีที่ใช้ในการแก้ไขปัญหาที่เหมาะสม เพื่อให้ได้คำตอบที่ต้องการ โดยการเลียนแบบพฤติกรรมการออกหาอาหารของผึ้ง ประชากรของขั้นตอนวิธีอาณานิคมผึ้งเทียมแบ่งออกเป็น 3 ประเภท คือ ผึ้งสำรวจ (Scout Bees) มีหน้าที่ในการสำรวจหาแหล่งอาหารแหล่งใหม่ ผึ้งงาน (Employed Bees) มีหน้าที่ออกหาแหล่งอาหารแล้วกลับมาให้ข้อมูลเกี่ยวกับแหล่งอาหารที่พบกับผึ้งเฝ้าดูที่รออยู่ที่รัง ผึ้งเฝ้าดู (Onlooker Bees) มีหน้าที่คำนวณค่าความเหมาะสม (Fitness) และเลือกแหล่งอาหารเพื่อส่งผึ้งงานไปเก็บอาหาร ถ้าแหล่งอาหารแหล่งใดถูกเก็บจนหมด ผึ้งงานที่เป็นเจ้าของแหล่งอาหารนั้นจะกลายเป็นผึ้งสำรวจ เพื่อค้นหาแหล่งอาหารแหล่งใหม่ โดยในธรรมชาติ ผึ้งจะมีการเปรียบเทียบคุณภาพของแหล่งอาหารที่ถาแผ่นดิน (Dance Floor) โดยทำการเต้นที่เรียกว่า “เต้นแกว่ง (Waggle Dance)” เพื่อแสดงข้อมูลเกี่ยวกับทิศทาง ระยะทางและปริมาณน้ำหวานของแหล่งอาหาร โดยผึ้งเฝ้าดูจะทำการเปรียบเทียบคุณภาพของแหล่งอาหารที่ดีที่สุด

หลายปีที่ผ่านมาได้มีนักวิจัยที่นำขั้นตอนวิธีเกี่ยวกับผึ้งไปประยุกต์ใช้ในการแก้ปัญหาต่างๆ เช่น ในปี 2012 S. Bitam [10] ได้เสนอการนำขั้นตอนวิธีวงจรผึ้ง (Bees Life Algorithm: BLA) ไปประยุกต์ใช้ในการจัดตารางการทำงานในการประมวลผลแบบกลุ่มเมฆ โดยเลียนแบบพฤติกรรมการสืบพันธุ์และการค้นหาแหล่งอาหารของผึ้ง เพื่อประเมินประสิทธิภาพการทำงานในแต่ละทรัพยากรของระบบการประมวลผลแบบกลุ่มเมฆ โดยมุ่งเน้นเพื่อลดเวลาในการทำงานและลดความซับซ้อนในการทำงานของระบบ และในปีเดียวกัน T. Mizan [11] ได้เสนอการจัดตารางงานในการประมวลผลแบบกลุ่มเมฆแบบผสม (Hybrid Cloud) โดยใช้ขั้นตอนวิธีวงจรผึ้งและขั้นตอนวิธีประเภทละโมภ (Greedy Algorithm) เพื่อลดเวลารวมในระบบให้น้อยลง

ดังนั้นผู้วิจัยจึงได้เสนอแนวคิดในการจัดตารางงานบนพื้นฐานของการกระจายงานแบบสมดุล และเวลารวมในการ

ทำงานน้อยที่สุด สำหรับการประมวลผลแบบกลุ่มเมฆ โดยเน้นในกรณีที่เครื่องคอมพิวเตอร์เสมือนมีสภาพแวดล้อมภายในเหมือนกันและต่างกัน โดยประยุกต์ใช้ขั้นตอนวิธีอาณานิคมผึ้งเทียมควบคู่กับการจัดลำดับงานแบบฮิวริสติก

3. วิธีการดำเนินงานวิจัย

การประมวลผลแบบกลุ่มเมฆจะมีการใช้ทรัพยากรที่แตกต่างกันตามความต้องการของผู้ใช้ในระบบ อีกทั้งการร้องขอเข้าใช้บริการต้องคำนึงถึงภาระงานของเครื่องแม่ข่ายที่จะให้บริการด้วย การจัดการงานเข้าใช้ทรัพยากรที่มีประสิทธิภาพจะส่งผลให้สามารถลดเวลารวมในการทำงานในระบบและส่งผลให้เกิดการกระจายงานแบบสมดุล ได้มีการนำขั้นตอนอาณานิคมผึ้งเทียมมาประยุกต์ใช้ในการจัดการงานบนพื้นฐานของการกระจายงานแบบสมดุลในระบบ และลดเวลารวมในการทำงานของระบบในการประมวลผลแบบกลุ่มเมฆ [12] โดยจุดมุ่งหมายของการวิจัยครั้งนี้ เพื่อหาขั้นตอนวิธีการจัดการงานให้เหมาะสมกับสภาพแวดล้อมภายในของเครื่องคอมพิวเตอร์เสมือน ทั้งในกรณีที่เครื่องคอมพิวเตอร์เสมือนมีสภาพแวดล้อมภายในเหมือนกันและต่างกัน และในการจัดการงานนั้นจะต้องอยู่บนพื้นฐานของการกระจายงานแบบสมดุลและเวลารวมในการทำงานน้อยที่สุดในระบบการประมวลผลแบบกลุ่มเมฆ

กำหนด $VM = \{VM_1, VM_2, VM_3, \dots, VM_N\}$; N คือ จำนวนของคอมพิวเตอร์เสมือน; $Task = \{task_1, task_2, task_3, \dots, task_K\}$; K คือ จำนวนของงานในระบบ รูปแบบการเข้าทำงานในระบบเป็นการทำงานที่ไม่สามารถจัดจังหวะหรือแทรกกลางกันได้ (Non-Preemptive Process) ขั้นตอนวิธีจัดการงานเพื่อเข้าใช้เครื่องคอมพิวเตอร์เสมือนโดยใช้ขั้นตอนอาณานิคมผึ้งเทียมควบคู่กับการจัดลำดับงานแบบฮิวริสติกแสดงดังภาพที่ 1

4. ผลการทดลอง

4.1 การตั้งค่าพารามิเตอร์สภาพแวดล้อมในการทดลอง

เครื่องมือที่ใช้ในสร้างแบบจำลองเพื่อประเมินประสิทธิภาพการทำงานของวิธีการที่นำเสนอ คือ CloudSim-3.0.1 การตั้งค่าของเครื่องคอมพิวเตอร์เสมือนแบ่งได้เป็น 2 กรณี คือ

กรณีที่ 1: ค่าพารามิเตอร์ของเครื่องคอมพิวเตอร์เสมือนที่มีสภาพแวดล้อมภายในเหมือนกัน คือ MIPS (Million Instructions Per Second is a measure of the processing speed of the

เริ่มต้น

- กำหนดค่าที่เหมาะสม (Fitness) เริ่มต้นให้กับแต่ละ VM_i ดังสมการที่ (1)

$$Fit_i = pe_num_i \times pe_mips_i + vm_bw_i \quad (1)$$

โดยที่ Fit_i = ค่า Fitness ของแต่ละ VM เครื่องที่ i

ทำซ้ำ

- ส่งผึ้งสังเกตการณ์ แบบสุ่มเข้าไปยัง VM_i และคำนวณค่า Fitness ดังสมการที่ (2) และ (3)

$$Fit_i = \frac{\sum_{j=1}^n task_length\ h_j}{Evaluate\ capacity\ of\ VM\ (Capa_i)} \quad (2)$$

โดยที่ $task_length$ = ขนาดของงาน

$$Capa_i = pe_num_i \times pe_mips_i + vm_bw_i \quad (3)$$

โดยที่ $Capa_i$ = ความสามารถในการทำงานของ VM เครื่องที่ i

- ผึ้งเฝ้าจะเลือก VM ที่มีค่า Fitness ดีที่สุดจำนวน m ตัว ซึ่ง m จะกลายเป็น Select Bee และ Site ที่ Select Bee ไปเยี่ยมชมมาแล้วจะถูกเรียกเป็นพื้นที่ที่ถูกเลือก (Neighborhood Search)
- กำหนดผึ้งจำนวนมากเข้าค้นหาในพื้นที่ที่ถูกเลือก ส่งผึ้งจำนวนมากกว่า e ที่เลือก จำนวนผึ้งที่ส่ง nep ตัว และคำนวณค่า Fitness ดังสมการที่ (4)

$$Fit_i = \frac{\sum_{j=1}^n task_length\ h_j + InputFile_length\ h}{Evaluate\ capacity\ of\ VM\ (Capa_i)} \quad (4)$$

โดยที่ $InputFile_length$ = ขนาดของงาน

- ผิ้นำข้อมูล Fitness กลับมาที่รังเพื่อทำการเลือก Site ที่ดีที่สุด
- กำหนดให้ งานจำนวน 1 งานเข้าทำงานใน VM ที่ถูกเลือก
- คำนวณเพื่อค่า Load balance ของระบบหลังจากมีการจัดงานให้เข้าทำงานใน VM จะมีการคำนวณภาระงานทั้งหมดในปัจจุบันของ VM ทั้งหมด โดยคำนวณจากค่าส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation: S.D.) เพื่อวัดภาระงานทั้งหมดใน VM ดังสมการที่ (5) - (7)

$$S.D. = \sqrt{\frac{1}{n} \sum_{j=0}^n (X_j - \bar{X})^2} \quad (5)$$

$$X_j = \frac{\sum_{i=1}^k task_length\ h_i}{capacity_j} \quad (6) \quad , \quad \bar{X} = \frac{\sum_{j=1}^n X_j}{n} \quad (7)$$

โดยที่ X_j = เวลารวมในการทำงานของแต่ละ VM สามารถคำนวณได้จากสมการที่ (6); \bar{X} = ค่าเฉลี่ยของเวลารวมในการทำงานของ VM ทั้งหมดในระบบสามารถคำนวณได้จากสมการที่ (7)

ถ้า $S.D \leq \bar{X}$ แสดงว่าระบบอยู่ในสภาวะสมดุล และ
 ถ้า $S.D > \bar{X}$ แสดงว่าระบบอยู่ในสภาวะไม่สมดุล

จนกระทั่ง (ทุกงานเข้าทำงานใน VM)

จบการทำงาน

ภาพที่ 1: รหัสเทียม (Pseudocode) แสดงขั้นตอนการจัดการงานในระบบการประมวลผลแบบกลุ่มเมฆ โดยใช้ขั้นตอนวิธีอาณานิคมผึ้งเทียม

computer) of PE = 2000, Number of PE per VM = 1, RAM = 1024, Bandwidth = 1000 Bit

กรณีที่ 2: ค่าพารามิเตอร์ของเครื่องคอมพิวเตอร์เสมือนที่มีสภาพแวดล้อมภายในต่างกัน คือ MIPS of PE = 1000- 2000,

Number of PE per VM = 1- 8, RAM = 512 - 6144,

Bandwidth = 1000 - 2000 Bit

การตั้งค่าพารามิเตอร์แบบจำลองเพื่อประเมินประสิทธิภาพของระบบการประมวลผลแบบกลุ่มเมฆแสดงดังตารางที่ 1

ตารางที่ 1: ค่าพารามิเตอร์ของแบบจำลองการประมวลผลแบบกลุ่มเมฆ

Type	Parameter	Value
Data Center	Number of Data Center	10
	Number of Host	5
	Type of Manager	Space_shared, Time_shared
	Number of PE ^a per Host	4
	Bandwidth	2000
	Host Memory (RAM) (MB ^b)	10240
	Data Center Cost (The cost of using processing in this resource)	10
Virtual Machine (VM)	Total number of VM	50
	Type of Manager	Time_shared
Task	Total number of Task	100 - 700
	Length of Task (MI ^c)	5000 - 20000
	Number of PE per requirement	1
	Type of Manager	10

PE^a = Processing Element, MB^b = Megabyte, MI^c is Million Instruction.

4.2 การตั้งค่าพารามิเตอร์ในขั้นตอนวิธีอาณานิคมผึ้งเทียม

การตั้งค่าพารามิเตอร์ที่ใช้ในส่วนของขั้นตอนวิธีอาณานิคมผึ้งเทียม แสดงค่าดังตารางที่ 2

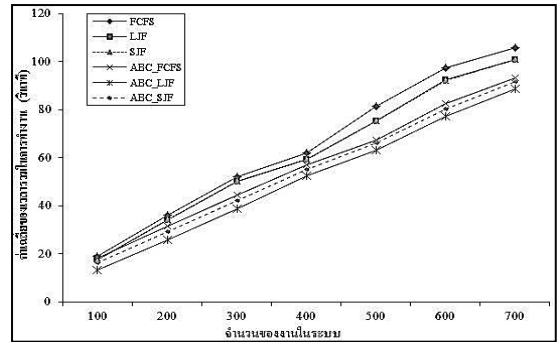
ตารางที่ 2: ค่าพารามิเตอร์ของขั้นตอนวิธีอาณานิคมผึ้งเทียม

Symbol	Parameter	Case I	Case II
n	Number of scout bees	2000	1500
m	Number of sites selected out of n visited site	5	10
e	Number of best site out of m select site	1	1
nep	Number of bees recruited for best e site	1600	1000
nsp	Number of bees recruited for other (m-e) selected sites	400	500

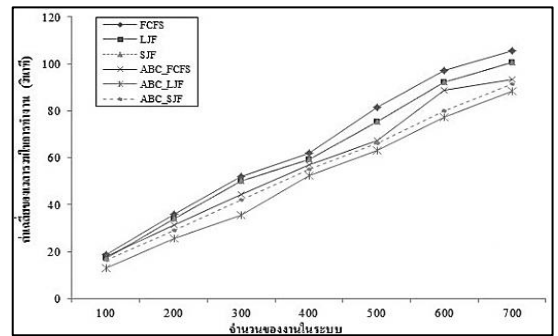
4.3 ผลการทดลอง

การประเมินประสิทธิภาพของวิธีที่นำเสนอ (ABC_FCFS, ABC_SJF และ ABC_LJF) จะทำการเปรียบเทียบกับการจัดลำดับงานแบบฮิวริสติก (FCFS, SJF และ LJF) โดยแบ่งการทดลองออกเป็น 2 กรณี คือ

กรณีที่ 1 : ทดสอบความสามารถในการทำงานเปรียบเทียบกับจำนวนของงานที่มีการเปลี่ยนแปลง กรณีเครื่องคอมพิวเตอร์เสมือนมีสภาพแวดล้อมภายในเหมือนกันทั้งหมด เมื่อจำนวนของ VM มีค่าคงที่ โดยแกน Y แสดงเวลาทำงานรวมเฉลี่ย (Average Processing Time) ของงานในระบบ



(a)

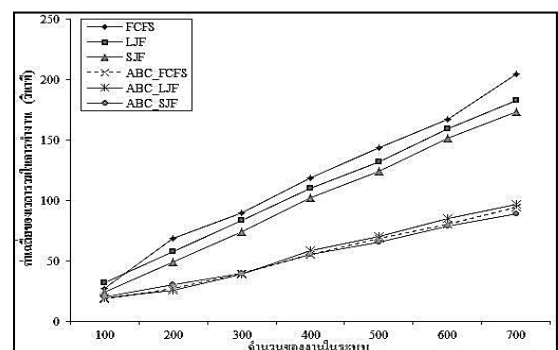


(b)

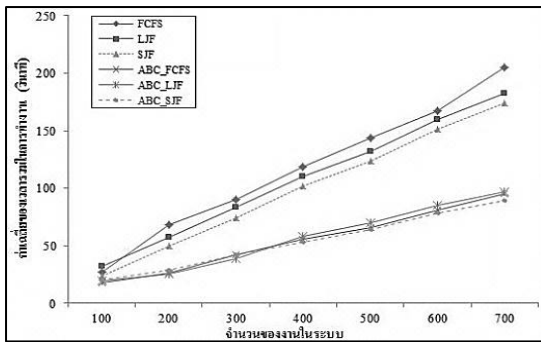
ภาพที่ 2: เวลาทำงานรวมเฉลี่ย เมื่องานมีปริมาณเพิ่มขึ้นและ VM มีสภาพแวดล้อมภายในเหมือนกัน โดยที่ (a) พารามิเตอร์ ABC Case I (b) พารามิเตอร์ ABC Case II

จากภาพที่ 2(a) และ 2(b) แสดงเวลาทำงานรวมเฉลี่ยโดยทำการเปรียบเทียบขั้นตอนการจัดตารางงานในเครื่องคอมพิวเตอร์เสมือนที่มีสภาพแวดล้อมภายในเหมือนกัน โดยใช้ขั้นตอนวิธีอาณานิคมผึ้งเทียมควบคู่กับการจัดลำดับงานแบบฮิวริสติกและการจัดลำดับงานแบบฮิวริสติกเพียงอย่างเดียว ผลจากการทดลองแสดงให้เห็นว่า ABC_LJF มีประสิทธิภาพในการจัดตารางงานดีที่สุด เมื่อปริมาณงานในระบบมีปริมาณเพิ่มมากขึ้น

กรณีที่ 2 : ทดสอบความสามารถในการทำงานเปรียบเทียบกับจำนวนของงานที่มีการเปลี่ยนแปลง กรณีเครื่องคอมพิวเตอร์เสมือนมีสภาพแวดล้อมภายในต่างกัน เมื่อ VM มีจำนวนคงที่ โดยแกน Y แสดงเวลาทำงานรวมเฉลี่ยของงานในระบบ



(a)



ภาพที่ 3: เวลาทำงานรวมเฉลี่ย เมื่องานมีปริมาณเพิ่มขึ้นและ VM มีสภาพแวดล้อมภายในต่างกัน โดยที่ (a) พารามิเตอร์ ABC Case I

(b) พารามิเตอร์ ABC Case II

จากภาพที่ 3(a) และ 3(b) แสดงเวลาทำงานรวมเฉลี่ยโดยทำการเปรียบเทียบขั้นตอนการจัดตารางงาน เพื่อเข้าใช้เครื่องคอมพิวเตอร์เสมือนที่มีสภาพแวดล้อมภายในต่างกัน โดยใช้ขั้นตอนวิธีอาณานิคมซึ่งเทียบควบคู่กับการจัดลำดับงานแบบิวริสติก และการจัดตารางแบบิวริสติกเพียงอย่างเดียว ผลจากการทดลองแสดงให้เห็นว่า ABC_SJF มีประสิทธิภาพในการจัดตารางการทำงานที่ดีที่สุด เมื่องานในระบบมีปริมาณมากกว่า 400 งานขึ้นไป

5. บทสรุป

งานวิจัยนี้ได้เสนอวิธีการจัดตารางงานบนพื้นฐานของการกระจายงานแบบสมดุลในการประมวลผลแบบกลุ่มเมฆโดยใช้ขั้นตอนวิธีอาณานิคมซึ่งเทียบควบคู่กับการจัดลำดับงานแบบิวริสติก โดยผลจากการทดลองแสดงให้เห็นว่า ในกรณีที่มีสภาพแวดล้อมภายในของเครื่องคอมพิวเตอร์เสมือนมีสภาพเหมือนกันควรใช้ขั้นตอนวิธีอาณานิคมซึ่งเทียบที่นำเสนอทำงานร่วมกับการจัดตารางงานโดยพิจารณาจากงานที่มีขนาดใหญ่ที่สุดเข้าทำงานก่อน (ABC_LJF) และในกรณีที่มีสภาพแวดล้อมภายในของเครื่องคอมพิวเตอร์เสมือนมีสภาพต่างกันควรใช้วิธีการที่นำเสนอควบคู่กับการจัดลำดับงานโดยพิจารณาจากงานที่มีขนาดเล็กที่สุดเข้าทำงานก่อน (ABC_SJF) เมื่อปริมาณของงานมีมากกว่า 400 งาน จะทำให้ประสิทธิภาพในการจัดตารางการทำงานเหมาะสมที่สุด ในการกระจายงานแบบสมดุลและเวลารวมในการทำงานน้อยที่สุด เมื่อปริมาณงานในระบบเพิ่มมากขึ้นและสภาพแวดล้อมภายในของระบบมีการเปลี่ยนแปลง

เอกสารอ้างอิง

- [1] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "A view of cloud computing," *Communications of the acm*, vol. 53, no. 4, pp. 50-58, 2010.
- [2] K. Dutta, "A smart job scheduling system for cloud computing service providers and users: modeling and simulation," in *1st Int'l Conf. on Recent Advances in Information Technology RAIT-2012*, 2012.
- [3] S. Sindhu, and S. Mukherjee, "Efficient task scheduling algorithms for cloud computing environment," *High Performance Architecture and Grid Computing, Communications in Computer and Information Science*, vol. 169, pp. 79-83, 2011.
- [4] D. Karaboga, and B. Basturk, "Artificial Bee Colony optimization algorithm for solving constrained optimization problems," *IFSA 2007*, LNAI 4529, pp. 789-798, 2007.
- [5] D. Karaboga, B. Gorkemli, C. Ozturk, and N. Karaboga, "A comprehensive survey: artificial bee colony (ABC) algorithm and applications," *Artificial Intelligence Review 2012, Springer Science Business Media B.V. 2012*, 2012.
- [6] Z.R.M. Azmi, K.A. Bakar, M.S. Shamsir, W.N.W. Manan, A.H. Abdullah, "Performance comparison of priority rule scheduling algorithms using different inter arrival time jobs in grid environment," *International Journal of Grid and Distributed Computing*, Vol. 4, No. 3, 61-70, 2011
- [7] Y. Fang, F. Wang, and J. Ge, "A task scheduling algorithm based on load balancing in cloud computing," in *Web Information Systems and Mining, Lecture Notes in Computer Science 2010*, Vol. 6318, pp. 271-277, 2010.
- [8] R. N. Calheiros, R. Ranjan, C. A. F. D. Rose, and R. Buyya, "CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms," *Software : Practice and Experience*, Vol. 41, No.1, pp. 23-50, Jan 2011.
- [9] B. Mondal, K. Dasgupta, and P. Dutta, "Load balancing in cloud computing using Stochastic Hill Climbing-A soft computing approach," in *Procedia Technology*, vol. 4, pp. 783-789, 2011.
- [10] S. Bitam, "Bees life algorithm for job scheduling in cloud computing," in *ICCIT 2012*, pp. 186-191, 2012.
- [11] T. Mizan, S. M. R. A. Masud, and R. Latip, "Modified bees life algorithm for job scheduling in hybrid cloud," in *IJET*, vol. 2, no.6, pp. 974-979, June 2012.
- [12] B. Kruekaew and W. Kimpan, "Virtual Machine Scheduling Management on Cloud Computing Using Artificial Bee Colony," *IMECS 2014*, HK, pp. 18-22, March 12-14, 2014.

วิธีการวางทับข้อมูลโดยการเปรียบเทียบมูลค่าสำหรับพื้นที่เก็บข้อมูลบนคลาวด์

Cost-Based Replacement Policy for Cloud Storage

สุเมธ สิริเรืองอำไพ (Sumet Siriruangapai)¹ และณัฐวุฒิ หนูไพโรจน์ (Natawut Nupairoj)²

ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์

จุฬาลงกรณ์มหาวิทยาลัย

¹Sumet.Si@Student.chula.ac.th, ²natawut.n@chula.ac.th

บทคัดย่อ

พื้นที่เก็บข้อมูลบนคลาวด์ถือเป็นหนึ่งในบริการที่ประสบความสำเร็จมากที่สุดของคลาวด์คอมพิวเตอร์ทั้งในแง่ของการใช้งานส่วนบุคคลและในระดับองค์กร โดยหนึ่งในปัจจัยสำคัญที่ส่งผลต่อความสำเร็จของพื้นที่เก็บข้อมูลบนคลาวด์ในระดับองค์กรนั้น ได้แก่ เรื่องค่าใช้จ่ายที่ต่ำกว่ามากหากเทียบกับการก่อตั้งหรือขยายขนาดศูนย์ข้อมูล (Data Center) ของตัวเอง อย่างไรก็ตามการคิดค่าใช้จ่ายของพื้นที่เก็บข้อมูลบนคลาวด์ในระดับองค์กร จำเป็นจะต้องพิจารณาอย่างรอบคอบเพราะค่าบริการที่เกิดขึ้นจากการถ่ายโอนข้อมูลอาจจะทำให้ค่าใช้จ่ายรวมมีมูลค่าสูงกว่าที่คาดการณ์ไว้ ซึ่งหนึ่งในวิธีที่มีการนำมาใช้แก้ไขปัญหานี้ก็คือ การนำแคชเข้ามาช่วยลดปริมาณการถ่ายโอนข้อมูลระหว่างผู้ใช้กับผู้ใช้บริการ แต่ในปัจจุบันวิธีการวางทับข้อมูลที่ถูกออกแบบมาสำหรับพื้นที่เก็บข้อมูลบนคลาวด์นั้น ยังมีไม่มากนัก

งานวิจัยนี้จะนำเสนอวิธีการวางทับข้อมูลแบบใหม่ ที่มีจุดประสงค์เพื่อลดค่าใช้จ่ายจากการถ่ายโอนข้อมูลและลดความล่าช้าของระบบ

คำสำคัญ: พื้นที่เก็บข้อมูลบนคลาวด์, ดาตาแคชซิง, วิธีการวางทับข้อมูล

Abstract

Cloud storage has become one of the most successful cloud computing services for personal and corporation use. One of the important factors contributing to the success of Cloud Storage at corporation level is much lower expense when compared to building or expanding their own data center. However, careful consideration

must be given when it comes to expense calculation at corporation level because data transfer fees may lead to higher expense than expected. One of the solutions that available for this problem is to use cache in order to reduce the traffic from transfer data between users and cloud providers but nowadays there are a few cache replacement policies that designed for cloud storage.

In this paper, we present a new cache replacement policy that aims to reduce the cost of data transfer as well as the latency and it has proved to work effectively.

Keyword: Cloud Storage, Data Caching, Cache Replacement Policy

1. บทนำ

พื้นที่เก็บข้อมูลบนคลาวด์ถือเป็นหนึ่งในบริการที่ประสบความสำเร็จมากที่สุดของคลาวด์คอมพิวเตอร์ทั้งในแง่ของการใช้งานส่วนบุคคลและในระดับองค์กร ดังจะเห็นได้จากบริการต่างๆ ที่ได้รับความนิยมในปัจจุบัน เช่น Dropbox [1] หรือ iCloud [2] ซึ่งสร้างความสะดวกสบายให้กับผู้ใช้งานเป็นอย่างมาก และในส่วนของระดับองค์กรเอง พื้นที่เก็บข้อมูลบนคลาวด์ก็เริ่มเป็นที่แพร่หลายมากขึ้น โดยสังเกตได้จากกลุ่มลูกค้าของ Amazon ที่มีทั้งมหาวิทยาลัยและบริษัทต่างๆ [3]

หนึ่งในปัจจัยสำคัญที่ทำให้พื้นที่เก็บข้อมูลบนคลาวด์ได้รับความนิยมขึ้นมาในระดับองค์กรนั้น เป็นเพราะมีค่าใช้จ่ายที่ถูกกว่าหากเทียบกับการสร้างหรือขยายขนาดศูนย์ข้อมูลของตัวเอง [4] แต่ในความเป็นจริงแล้วการคิดค่าใช้จ่ายของพื้นที่เก็บข้อมูลบนคลาวด์ในระดับองค์กรจะมีความซับซ้อนมากกว่าแบบส่วนบุคคลที่เสียค่าใช้จ่ายเฉพาะค่าเช่าพื้นที่ เพราะจะมีเรื่องของ

ค่าใช้จ่ายที่เกิดขึ้นจากการถ่ายโอนข้อมูลเข้ามาเกี่ยวข้อง ซึ่งอาจทำให้ค่าใช้จ่ายขององค์กรที่มีการถ่ายโอนข้อมูลในปริมาณมาก สูงกว่าที่คาดการณ์ไว้

ในงานวิจัย [5] เราได้นำเสนอการนำคลาวด์แคชซึ่งมาใช้เพื่อลดค่าใช้จ่ายรวมของความเป็นเจ้าของบนพื้นที่เก็บข้อมูลบนคลาวด์และเพื่อเพิ่มประสิทธิภาพให้แก่ระบบ ซึ่งผลการทดลองก็แสดงให้เห็นว่าคลาวด์แคชซึ่งมีวิธีการวางทับข้อมูลแบบ LRU สามารถลดค่าใช้จ่ายรวมของความเป็นเจ้าของและความล่าช้าของระบบลงได้ ซึ่งในงานวิจัยนี้เราจะนำเสนอวิธีการวางทับข้อมูลแบบใหม่ ที่มีจุดประสงค์เพื่อลดค่าใช้จ่ายจากการถ่ายโอนข้อมูล โดยเราจะใช้การะงาน (Trace File) ที่เราเก็บข้อมูลมาจาก NFS Server ของคณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัยมาจำลองการทำงาน ซึ่งผลที่ออกมาแสดงให้เห็นว่าวิธีการวางทับข้อมูลของเราทำงานได้อย่างมีประสิทธิภาพ

2. งานวิจัยที่เกี่ยวข้อง

2.1 การถ่ายโอนข้อมูลในพื้นที่เก็บข้อมูลบนคลาวด์

แม้ว่าพื้นที่เก็บข้อมูลบนคลาวด์จะใช้เงินลงทุนที่ต่ำมากเมื่อเทียบกับศูนย์ข้อมูล แต่งานวิจัยที่ผ่านมาได้ชี้ให้เห็นว่าค่าบริการที่เกิดขึ้นจากการถ่ายโอนข้อมูลอาจจะทำให้ค่าใช้จ่ายรวมมีมูลค่าที่สูงขึ้นอย่างมีนัยสำคัญ จากงานวิจัย [9] ที่สร้างแบบจำลองทางสมการขึ้นมาเพื่อใช้ในการคำนวณว่าองค์กรนั้นๆ เหมาะสมที่จะใช้บริการคลาวด์คอมพิวติ้งหรือไม่ จะพบว่าหนึ่งในตัวแปรหลักของสมการก็คือปริมาณข้อมูลที่มีการถ่ายโอน ซึ่งหากเป็นองค์กรที่มีลักษณะการถ่ายโอนข้อมูลบ่อยหรือมีปริมาณสูง จะทำให้ความเหมาะสมของคลาวด์คอมพิวติ้งกับองค์กรแปรผกผันกับปริมาณข้อมูลที่มีการถ่ายโอน

งานวิจัยของเรา [5] ได้นำเสนอผลการเปรียบเทียบค่าใช้จ่ายรวมของความเป็นเจ้าของระหว่างพื้นที่เก็บข้อมูลท้องถิ่น (Local Storage) กับพื้นที่เก็บข้อมูลบนคลาวด์ ซึ่งในแง่ของค่าใช้จ่ายที่เกิดขึ้นจากพื้นที่เก็บข้อมูล ระบบคลาวด์จะมีค่าใช้จ่ายที่ต่ำกว่าพื้นที่เก็บข้อมูลท้องถิ่น แต่เมื่อมีการรวมค่าบริการจากการถ่ายโอนข้อมูลเข้าไป ค่าใช้จ่ายรวมของความเป็นเจ้าของในระบบคลาวด์จะมีมูลค่าที่สูงกว่า และเมื่อพิจารณาค่าใช้จ่ายรวมของความเป็นเจ้าของในระบบที่มีการใช้งานแคช ค่าใช้จ่ายจะมีแปรแปรผกผันกับขนาดของแคช

ปัจจุบันวิธีการแก้ไขปัญหานี้ที่ถูกนำเสนอในงานวิจัย [5, 8, 10] คือการนำแคชเข้ามาใช้งานเพื่อลดปริมาณการถ่ายโอนข้อมูลระหว่างผู้ให้บริการ ซึ่งผลที่ได้ปรากฏว่าแคชเป็นตัวเลือกที่มีประสิทธิภาพและเหมาะสมกับปัญหานี้

2.2 วิธีการวางทับข้อมูลสำหรับพื้นที่เก็บข้อมูลบนคลาวด์

จากงานวิจัย [6, 7] ได้มีการชี้ให้เห็นว่า วิธีการวางทับข้อมูลแบบ LRU [11] ซึ่งถือได้ว่าเป็นวิธีการวางทับข้อมูลที่มีประสิทธิภาพมากที่สุดในหลายๆระบบ กลับไม่เหมาะสมสำหรับพื้นที่เก็บข้อมูลบนเครือข่าย (Network Storage) อย่างไรก็ตามวิธีการวางทับข้อมูล LFU [11] สามารถทำงานได้ดีกว่าในสภาพแวดล้อมเช่นนี้ เนื่องจากทั้งเครื่องลูกข่ายและเครื่องแม่ข่ายต่างก็มีแคชเป็นของตัวเอง ทำให้การเรียกใช้ข้อมูลจากแคชของพื้นที่เก็บข้อมูลบนเครือข่ายไม่มีพฤติกรรมในรูปแบบของ Recency-Based ดังนั้นเราจะนำเอาวิธีการวางทับข้อมูล LFU มาประยุกต์ใช้ในการออกแบบวิธีการวางทับข้อมูลของเรา

Banditwattanawong [10] ได้นำเสนอวิธีการวางทับข้อมูลสำหรับคลาวด์โดยมีจุดประสงค์เพื่อปรับปรุงประสิทธิภาพและลดค่าใช้จ่ายที่เกิดจากการถ่ายโอนข้อมูล ซึ่งผลการทดสอบได้แสดงให้เห็นว่าวิธีการวางทับข้อมูลของเขามีประสิทธิภาพที่ดีกว่าวิธีอื่นๆ อย่างไรก็ตามการะงานที่ถูกนำมาใช้ในงานวิจัย [10] ไม่ใช่การะงานที่สะท้อนถึงสภาพแวดล้อมที่แท้จริงของการทำงานภายใต้พื้นที่เก็บข้อมูลบนคลาวด์ เนื่องจากการะงานทั้งหมดเป็นการะงานที่ได้มาจากเว็บแคช และตัวแปรบางส่วนในสมการที่ใช้คำนวณมูลค่าข้อมูลยังได้รับอิทธิพลมาจากลักษณะการะงานประเภทเว็บแคชอีกด้วย เช่น ค่าความหน่วง (Latency) นอกจากนี้ในการหาค่า Time-To-Live เพื่อนำมาใช้กับสมการยังจำเป็นต้องประมวลผลการะงานก่อนที่จะนำมาทดลอง ซึ่งนั่นจะทำให้เกิดความยุ่งยากในกรณีที่ต้องการนำไปปรับใช้งานจริง

$$s_i \times c_i \times l_i \times f_i \times TTL_i \quad (1)$$

สมการที่ 1 เป็นสมการคำนวณมูลค่าข้อมูลในงานวิจัยของ Banditwattanawong [10] ซึ่ง s_i คือขนาดของข้อมูล i c_i คือค่าใช้จ่ายในการดาวน์โหลดข้อมูล i l_i คือความหน่วงของการดาวน์โหลดข้อมูล i f_i คือความถี่ในการเรียกใช้งานข้อมูล i และ TTL_i คือค่า Time-To-Live ของข้อมูล i

ดังนั้นหากเราตัดตัวแปร l_i และ TTL_i ซึ่งเป็นตัวแปรที่ไม่ปรากฏอยู่ในภาระงานจาก NFS Server ของเราออกจากสมการ จะส่งผลให้วิธีการวางทับข้อมูลตามสมการที่ 1 ทำงานเหมือนกับ LFU ซึ่งจะประสบปัญหาแชนทำงานผิดพลาดจากข้อมูลที่เคยได้รับความนิยมน้อยมากในอดีต [11] ฉะนั้นเราจะใช้วิธีการวางทับข้อมูล LFU เป็นตัวแทนอัลกอริทึมของ Banditwattanawong ในการทดลองของเรา ทั้งนี้วิธีการจัดเก็บข้อมูลลงแชนของเราจะแบ่งข้อมูลออกเป็นบล็อก ทำให้บล็อกข้อมูลทั้งหมดมีขนาดเท่ากัน จึงทำให้ Latency ของข้อมูลทุกตัวมีค่าเท่ากัน

3. วิธีการวางทับข้อมูลโดยการเปรียบเทียบมูลค่า

ในส่วนนี้จะอธิบายถึงวิธีการวางทับข้อมูลของเรานี้ที่มีชื่อว่า Cost-Based Replacement Policy ซึ่งหลังจากนี้เราจะใช้ตัวย่อว่า CBRP

3.1 อัลกอริทึม

เมื่อผู้ใช้ร้องขอข้อมูล ระบบจะค้นหาข้อมูลที่อยู่ภายในแชน หากพบก็จะส่งข้อมูลนั้นๆกลับไปยังผู้ใช้ แล้วทำการปรับปรุง (Update) มูลค่าของข้อมูลนั้นๆที่อยู่ในแชน แต่หากไม่พบระบบจะทำการร้องขอไปที่คลาวด์ เมื่อได้ข้อมูลแล้วก็จะส่งข้อมูลกลับไปให้ผู้ใช้ แล้วจึงกำหนดมูลค่าให้แก่ข้อมูลนั้นๆ ก่อนเก็บลงในแชน แต่ถ้าแชนมีที่ว่างไม่เพียงพอระบบจะค้นหาข้อมูลที่มีมูลค่าน้อยที่สุด แล้วตรวจสอบว่าข้อมูลนั้นๆมีการเปลี่ยนแปลงไปจากต้นฉบับหรือไม่ ถ้าไม่มีการเปลี่ยนแปลงระบบจะนำข้อมูลนั้นๆออกไป แล้วบันทึกข้อมูลใหม่ลงแชน แต่ถ้าข้อมูลมีการเปลี่ยนแปลงไป ระบบจะอัปเดตค่าข้อมูลนั้นๆไปที่คลาวด์ก่อนนำออกจากแชน แล้วจึงบันทึกข้อมูลใหม่ลง

ด้านการเขียนข้อมูลเราจะใช้วิธีการที่เรียกว่า “Write-back” โดยเมื่อผู้ใช้เขียนข้อมูล ระบบจะค้นหาข้อมูลนั้นๆภายในแชน หากพบก็จะปรับปรุงข้อมูลให้ตรงกับข้อมูลของผู้ใช้ แล้วทำการปรับปรุงมูลค่าของข้อมูลนั้นๆที่อยู่ในแชน จากนั้นจึงบันทึกไว้ว่าข้อมูลมีการเปลี่ยนแปลงไปจากต้นฉบับ หากไม่พบระบบจะทำการตรวจสอบพื้นที่ในแชนเพื่อดูว่ายังมีพื้นที่ว่างเหลืออยู่หรือไม่ หากมีก็จะทำการเขียนข้อมูลลงไปในแชน แล้วทำการกำหนดมูลค่าให้กับข้อมูลนั้นๆ แล้วจึงบันทึกว่าข้อมูลนั้นๆมีการเปลี่ยนแปลงไปจากต้นฉบับ แต่ถ้ามีที่ว่างไม่เพียงพอ

ระบบก็จะจัดการนำข้อมูลที่มีมูลค่าน้อยที่สุดออกไป โดยใช้วิธีเดียวกันกับการร้องขอข้อมูลในข้างต้น

```

algorithm CACHING
input DATA REQUEST
begin
    if REQUESTED DATA found in CACHE
        update DATA WEIGHT by using DATA WEIGHT EQUATION
        Return REQUESTED DATA to USER
    else
        download REQUESTED DATA from CLOUD STORAGE
        assign DATA WEIGHT by using DATA WEIGHT EQUATION
        Return REQUESTED DATA to USER
    if CACHE has free space
        save REQUESTED DATA into CACHE
    else
        if LEAST WEIGHT DATA modified
            write LEAST WEIGHT DATA to CLOUD STORAGE
            evict LEAST WEIGHT DATA from CACHE
            save REQUESTED DATA into CACHE
        else
            evict LEAST WEIGHT DATA from CACHE
            save REQUESTED DATA into CACHE
end

```

```

algorithm WRITE BACK POLICY
input DATA WRITING
begin
    if WRITTEN DATA found in CACHE
        update DATA WEIGHT by using DATA WEIGHT EQUATION
        update DATA in CACHE
    else
        assign DATA WEIGHT by using DATA WEIGHT EQUATION
    if CACHE has free space
        write DATA into CACHE
    else
        if LEAST WEIGHT DATA modified
            write LEAST WEIGHT DATA to CLOUD STORAGE
            evict LEAST WEIGHT DATA from CACHE
            write DATA into CACHE
        else
            evict LEAST WEIGHT DATA from CACHE
            write DATA into CACHE
        mark DATA as modified
end

```

ภาพที่ 1: การทำแชนซิงและเขียนข้อมูล “Write-back”

3.2 การกำหนดมูลค่าให้กับข้อมูล

ตามที่ได้กล่าวไปแล้วในข้างต้นว่า LFU สามารถทำงานได้มีประสิทธิภาพมากกว่า LRU เมื่อถูกนำมาใช้งานกับพื้นที่เก็บ

ข้อมูลบนเครือข่าย ดังนั้นเราจะนำ LFU มาประยุกต์ใช้งานโดยการเพิ่มขนาดของข้อมูลและค่าบริการจากการถ่ายโอนข้อมูลเข้าไป ซึ่งปัญหาหลักของ LFU คือการที่ข้อมูลที่เคยถูกเรียกใช้งานหลายๆ อาจจะถูกทิ้งในแคชตลอดไป ซึ่ง 1 ในวิธีการแก้ปัญหาที่ถูกพัฒนาขึ้นมาได้แก่ LFU with Dynamic Aging policy (LFU-DA) [13] ซึ่งไม่เพียงแต่จะแก้ปัญหาได้เท่านั้น LFU-DA ยังมีความซ้ำซ้อนน้อยกว่า LFU-Aging policy [14] อีกด้วย ผู้ใช้ไม่จำเป็นต้องกำหนดค่าตัวแปร เพื่อให้สมการทำงานได้เหมาะสมกับภาระงานนั้นๆ โดยสมการที่เราใช้ในการคำนวณมูลค่าของข้อมูลจะเป็นไปตามสมการที่ 2

$$w_i = (f_i \times c_i \times s_i) + w_{evict} \quad (2)$$

ซึ่ง f_i คือความถี่ในการเรียกใช้งานข้อมูล i c_i คือค่าใช้จ่ายในการถ่ายโอนข้อมูล i s_i คือขนาดของข้อมูล i และ w_{evict} คือมูลค่าของข้อมูลล่าสุดที่ถูกนำออกไปจากแคช ซึ่งมีหน้าที่แก้ปัญหาที่มีลักษณะดังที่กล่าวไว้ในข้างต้น โดย w_{evict} จะมีค่าเริ่มต้นเป็น 0

4. การทดลอง

4.1 ภาระงาน

ในการทดลองของเราจะใช้ภาระงานที่เราเก็บข้อมูลมาจาก NFS Server ขนาด 4 เทราไบต์ที่อยู่ในคณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ตั้งแต่วันที่ 28 ส.ค. 2554 จนถึงวันที่ 3 กันยายน 2554 ซึ่งจะมีคุณลักษณะของภาระงานตามที่แสดงอยู่ในตารางที่ 1 ทั้งนี้ HR_∞ และ BHR_∞ คือค่า Hit Rate และ Byte Hit Rate เมื่อแคชมีเนื้อที่ไม่จำกัด และเรากำหนดให้แบบจำลองของเราแบ่งข้อมูลออกเป็นบล็อกที่มีขนาด 4,096 ไบต์ จึงทำให้ HR_∞ และ BHR_∞ มีค่าเท่ากัน

ตารางที่ 1: คุณลักษณะของภาระงาน

การร้องขอข้อมูลทั้งหมด	244,802,026 ครั้ง (933.85 GBs)
การร้องขอข้อมูลที่ไม่ซ้ำกัน	125,801,287 ครั้ง (479.89 GBs)
HR_∞	ร้อยละ 51.39
BHR_∞	ร้อยละ 51.39

4.2 การวัดประสิทธิภาพ

เราจะวัดประสิทธิภาพโดยใช้ค่า ปริมาณการถ่ายโอนข้อมูล ปริมาณแบนด์วิดท์ที่ลดลง Hit Rate Byte Hit Rate ค่าใช้จ่าย

ค่าความหน่วง ผลกระทบจากค่าบริการที่เปลี่ยนแปลง และค่าบริการเมื่อใช้งานผู้ใช้ให้บริการคลาวด์หลายแห่ง และเนื่องจากการวางทับข้อมูลแบบ LRU เป็นวิธีที่ถูกใช้งานอย่างแพร่หลายในระบบต่างๆ เราจึงนำ LRU มาเปรียบเทียบกับประสิทธิภาพของ CBRP ด้วย

งานวิจัยนี้มีจุดประสงค์หลักคือการลดค่าใช้จ่ายที่เกิดขึ้นจากการถ่ายโอนข้อมูล เพื่อที่จะทำให้พื้นที่เก็บข้อมูลบนคลาวด์มีค่าใช้จ่ายที่ถูกกว่าพื้นที่เก็บข้อมูลท้องถิ่น เราจะกำหนดขนาดของแคชให้อยู่ที่ร้อยละ 10-50 ของจำนวนข้อมูลที่ถูกร้องขอโดยไม่ซ้ำกัน และกำหนดให้บล็อกข้อมูลมีขนาด 4,096 และ 8,192 ไบต์ ยกเว้นการวัดค่าผลกระทบจากค่าบริการที่เปลี่ยนแปลง และค่าบริการเมื่อใช้งานผู้ใช้ให้บริการคลาวด์หลายแห่ง ที่จะกำหนดให้แคชมีขนาดร้อยละ 20 ของจำนวนข้อมูลที่ถูกร้องขอโดยไม่ซ้ำกัน และบล็อกข้อมูลมีขนาด 4,096

4.3 แบบจำลองค่าบริการ

ในงานวิจัยนี้เราจะใช้อัตราค่าบริการจาก IBM SmartCloud Enterprise [12] ซึ่งมีข้อมูลโดยสรุปดังตารางที่ 2

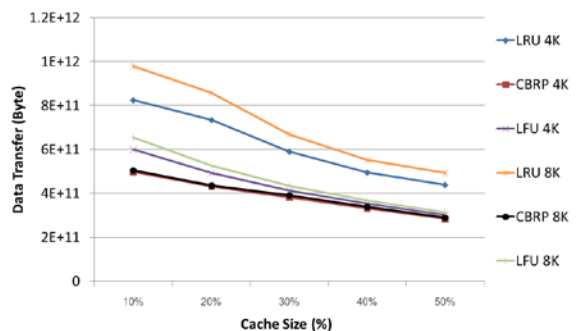
ตารางที่ 2: อัตราค่าบริการของ IBM SmartCloud Enterprise

ชนิดของบริการ	ค่าบริการ
Upload traffic	\$0.15 GB
Download traffic	\$0.15 GB

4.4 ผลการทดลอง

ในส่วนนี้เราจะเปรียบเทียบผลการทดลองระหว่าง CBRP LRU และ LFU

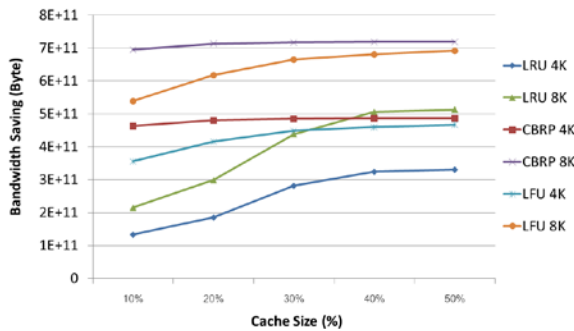
Data Transfer



ภาพที่ 2: ปริมาณการถ่ายโอนข้อมูล

จากภาพที่ 2 จะเห็นว่า การถ่ายโอนข้อมูลของ CBRP มีปริมาณที่น้อยกว่า LRU และ LFU ในแคชและบล็อกข้อมูลทุกขนาด และ LRU มีปริมาณการถ่ายโอนข้อมูลมากที่สุด

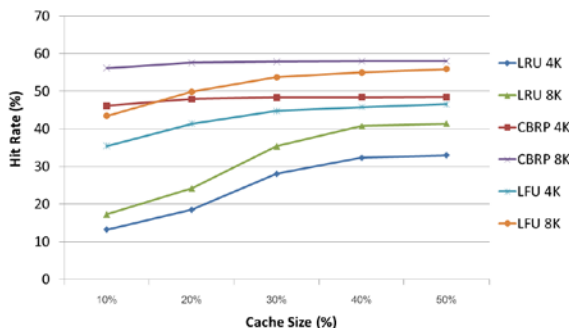
Bandwidth Saving



ภาพที่ 3: ปริมาณแบนด์วิดท์ที่ลดลง

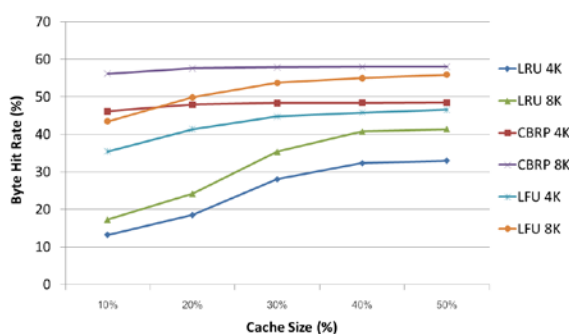
จากภาพที่ 3 จะเห็นว่า CBRP สามารถลดปริมาณแบนด์วิดท์ลงได้มากที่สุด เมื่อเทียบกับ LFU และ LRU ในกรณีที่บล็อกข้อมูลมีขนาดเท่ากัน

Hit Rate



ภาพที่ 4: Hit Rate

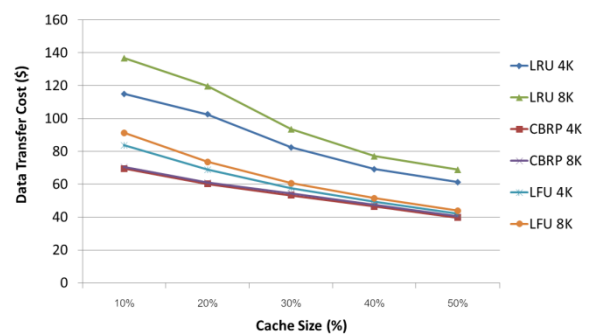
Byte Hit Rate



ภาพที่ 5: Byte Hit Rate

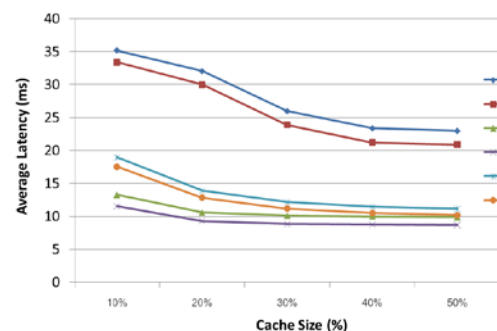
จากภาพที่ 4 และ 5 จะสังเกตเห็นว่ากราฟผลการทดลองที่ออกมา มีลักษณะเหมือนกัน ซึ่งเป็นผลมาจากการแบ่งข้อมูลออกเป็นบล็อกนั่นเอง โดยผลที่ออกมาจะพบว่า CBRP มีประสิทธิภาพที่เหนือกว่าทั้ง LFU และ LRU และประสิทธิภาพจะเริ่มคงที่เมื่อแคชมีขนาดร้อยละ 20 ในขณะที่ LFU และ LRU จะต้องรอจนกว่าแคชจะมีขนาดที่ร้อยละ 40 ซึ่งนั่นแสดงให้เห็นว่า CBRP สามารถทำงานได้อย่างมีประสิทธิภาพแม้แคชจะมีขนาดที่ไม่ใหญ่มากก็ตาม

Cost



ภาพที่ 6: ค่าใช้จ่าย

Latency

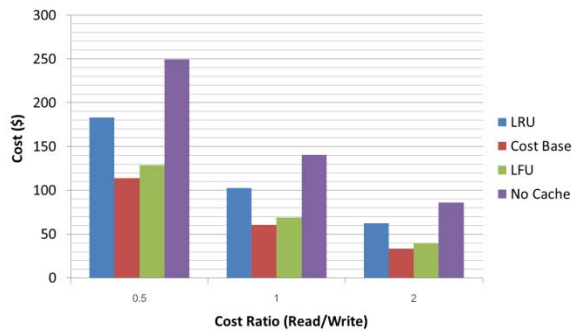


ภาพที่ 7: ค่าความหน่วง

ในส่วนของการค่าใช้จ่ายและค่าความหน่วงที่เกิดขึ้น ดังภาพที่ 7 และ 8 จะเห็นได้ว่า CBRP ก็ยังคงมีประสิทธิภาพที่สูงกว่า LRU และ LFU โดยที่ CBRP กับ LFU จะมีประสิทธิภาพที่ใกล้เคียงกันมาก เนื่องจาก CBRP มีพื้นฐานที่มาจาก LFU และเป็นการทดลองที่ใช้งานพื้นที่เก็บข้อมูลบนคลาวด์จากผู้ให้บริการเพียงแห่งเดียว นอกจากนี้บล็อกข้อมูลยังมีขนาดเท่ากันอีกด้วย จึงทำให้ค่า

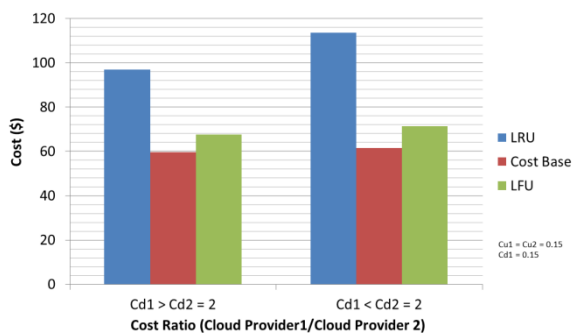
ของ c_i และ s_i ในสมการที่ 2 ไม่ส่งผลกระทบต่อผลลัพธ์มากนัก

Effect of Cost Ratio



ภาพที่ 8: ผลกระทบจากค่าบริการที่เปลี่ยนไป

Many Cloud Provider



ภาพที่ 9: ค่าบริการเมื่อใช้งานผู้ให้บริการคลาวด์หลายแห่ง

ภาพที่ 8 เป็นการนำเสนอผลกระทบจากค่าบริการที่เปลี่ยนไป โดยจะมีการเปลี่ยนอัตราค่าบริการในส่วนของการอัปโหลดเป็น 0.075 0.15 และ 0.3 \$/GB ตามลำดับ และคงค่าบริการการดาวน์โหลดไว้ที่ 0.15 \$/GB โดยผลที่ออกมาแสดงให้เห็นว่าการอัปโหลดข้อมูลส่งผลกระทบต่อค่าบริการสูงกว่าที่เราคาดการณ์ไว้ภายใต้ภาระงานนี้ ซึ่งนั่นยังทำให้เราเห็นถึงประโยชน์จากการที่ในปัจจุบันผู้ให้บริการคลาวด์หลายแห่งได้ยกเลิกการเก็บค่าบริการในส่วนนี้ไปแล้ว

ส่วนสุดท้ายที่เราให้ความสำคัญกับงานวิจัยนี้ได้แก่เรื่องการเปรียบเทียบค่าใช้จ่ายระหว่าง LRU และ CBRP เมื่อมีการใช้บริการพื้นที่เก็บข้อมูลบนคลาวด์จากผู้ให้บริการมากกว่า 1 แห่ง ซึ่งจากภาพที่ 9 จะเห็นว่าเมื่อค่าบริการในส่วนของการดาวน์โหลดจากผู้ให้บริการหมายเลข 2 ถูกกว่าผู้ให้บริการหมายเลข 1 สองเท่า ค่าบริการจากการถ่ายโอนข้อมูลของ CBRP จะอยู่ที่

59.5 \$ ในขณะที่ LRU จะมีมูลค่าสูงถึง 96.91 \$ และ LFU จะอยู่ที่ 67.49 \$ และเมื่อค่าบริการในส่วนของการดาวน์โหลดจากผู้ให้บริการหมายเลข 2 แพงกว่าผู้ให้บริการหมายเลข 1 สองเท่า ค่าบริการจากการถ่ายโอนข้อมูลของ CBRP จะอยู่ที่ 61.5 \$ ในขณะที่ LRU จะมีมูลค่า 113.51 \$ และ LFU จะอยู่ที่ 71.3 \$ ซึ่งจากทั้ง 2 กรณีจะพบว่า CBRP ทำให้เกิดค่าบริการที่น้อยที่สุด นอกจากนี้ยังมีส่วนต่างของค่าใช้จ่ายบริการจากทั้ง 2 กรณีอยู่ที่ 2 \$ เท่านั้นหรือคิดเป็น ร้อยละ 3.36 ของค่าบริการเดิม ในขณะที่ LRU กลับมีส่วนต่างที่มากถึง 16.6 \$ หรือคิดเป็น ร้อยละ 17.13 ของค่าบริการเดิม และสำหรับ LFU นั้นส่วนต่างของราคาจะอยู่ที่ 3.81 \$ หรือคิดเป็นร้อยละ 5.65 ของค่าบริการเดิม ซึ่งเราสามารถสรุปได้ว่า CBRP สามารถจัดการกับเรื่องค่าใช้จ่ายได้มีประสิทธิภาพมากที่สุด แม้ว่าจะมีการเปลี่ยนแปลงอัตราค่าบริการก็ตาม ทั้งนี้ในการทดลองเราได้กำหนดค่าบริการในการอัปโหลดของผู้ให้บริการทั้ง 2 แห่งเป็น 0.15 \$/GB และได้กำหนดค่าบริการในส่วนของการดาวน์โหลดจากผู้ให้บริการหมายเลข 1 เป็น 0.15 \$/GB ในส่วนการเลือกพื้นที่เก็บข้อมูลจากผู้ให้บริการทั้งสองแก่ข้อมูลต่างๆ เราจะใช้วิธีนำ File ID มาหาค่าเศษเหลือ (Modulus operator) เพื่อดูว่าจะเก็บข้อมูลนั้นๆ ไว้ที่ใคร

5. สรุป

บ่อยครั้งที่ค่าบริการที่เกิดจากการถ่ายโอนข้อมูลในพื้นที่เก็บข้อมูลบนคลาวด์มักถูกมองข้ามไป ทั้งที่ความเป็นจริงแล้วค่าบริการในส่วนนี้ส่งผลกระทบต่ออย่างมากกับค่าใช้จ่ายโดยรวม ซึ่งนั่นอาจจะทำให้ค่าใช้จ่ายของพื้นที่เก็บข้อมูลบนคลาวด์มีมูลค่าที่สูงกว่าพื้นที่เก็บข้อมูลท้องถิ่น

งานวิจัยนี้มีจุดประสงค์เพื่อลดปริมาณการถ่ายโอนข้อมูลลง โดยการนำเสนอวิธีการวางทับข้อมูลแบบใหม่ที่มีชื่อว่า Cost-Based Replacement Policy ซึ่งผลการทดลองที่ออกมาได้แสดงให้เห็นว่าวิธีการวางทับข้อมูลของเรามีประสิทธิภาพที่เหนือกว่าในทุกๆ ด้าน และน่าจะเป็นประโยชน์กับองค์กรที่ต้องการนำพื้นที่เก็บข้อมูลบนคลาวด์มาประยุกต์ใช้ระบบที่มีอยู่แล้วทั้งในแง่ของค่าใช้จ่ายและประสิทธิภาพ

เอกสารอ้างอิง

- [1] <http://www.dropbox.com/>
- [2] <http://www.apple.com/icloud/>
- [3] <http://aws.amazon.com/solutions/case-studies/>
- [4] Michael Armbrust, Armando Fox, Rean Griffith, Anthony D. Joseph, Randy Katz, Andy Konwinski, Gunho Lee, David Patterson, Ariel Rabkin, Ion Stoica, and Matei Zaharia. "Above the Clouds: A Berkeley View of Cloud Computing". February 2009.
- [5] Sumet Siriruangampai and Natawut Nupairoj. "Lower Storage TCO with Local Cloud Caching." *In Proceedings of the 2011 Thai Grid and Cloud Conference (TGCC '11)*, Bangkok, Thailand, June 2011, pp. 22-26.
- [6] Willick, D. L.; Eager, D.L.; Bunt, R.B., "Disk cache replacement policies for network file servers" *Distributed Computing Systems, 1993., Proceedings the 13th International Conference*, May 1993, pp.2-11.
- [7] Yusuke Nagasako and Saneyasu Yamaguchi. "A Server Cache Size Aware Cache Replacement Algorithm for Block Level Network Storage". *In Proceedings of the 2011 Tenth International Symposium on Autonomous Decentralized Systems (ISADS '11)*. IEEE Computer Society, Washington, DC, USA, pp. 573-576.
- [8] Mayur R. Palankar, Adriana Iamnitchi, Matei Ripeanu, and Simson Garfinkel. "Amazon S3 for science grids: a viable solution?". *In Proceedings of the 2008 international workshop on Data-aware distributed computing (DADC '08)*. ACM, New York, NY, USA, 55-64.
- [9] Subhas Chandra Misra, Arka Mondal, "Identification of a company's suitability for the adoption of cloud computing and modelling its corresponding Return on Investment", *Mathematical and Computer Modelling, Volume 53*, Issues 3-4, February 2011, pp. 504-521.
- [10] Thepparit Banditwattanawong. "From web cache to cloud cache". *In Proceedings of the 7th international conference on Advances in Grid and Pervasive Computing (GPC'12)*, Ruixuan Li, Jiannong Cao, and Julien Bourgeois (Eds.). Springer-Verlag, Berlin, Heidelberg, pp. 1-15.
- [11] Stefan Podlipnig and Laszlo Böszörményi. 2003. "A survey of Web cache replacement strategies". *ACM Comput. Surv.* 35, December 2003.
- [12] http://www-935.ibm.com/services/us/en/cloud-enterprise/contracts/charges_schedule.html
- [13] Martin Arlitt, Ludmila Cherkasova, John Dilley, Rich Friedrich, and Tai Jin. 2000. "Evaluating content management techniques for Web proxy caches". *SIGMETRICS Perform. Eval. Rev.* 27, March 2000, pp.3-11.
- [14] J. Robinson and M. Devarakonda, "Data Cache Management Using Frequency-Based Replacement", *Proceedings of the 1990 ACM SIGMETRICS Conference on the Measurement and Modeling of Computer Systems*, Boulder, CO, May 1990, pp. 134-142.

ระบบจัดการอัปโหลดไฟล์ผ่านโปรโตคอลเว็บซ็อกเก็ตสำหรับแอนดรอยด์สมาร์ทโฟน

File Upload Manager via Websocket Protocol for Android Smartphone

วรพงษ์ รัตน์ประสิทธิ์ (Worapong Rataprasit)¹ และ เชษฐ พัฒโนทัย (Chate Patanothai)²

¹ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

²ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

Worapong.Ra@student.chula.ac.th, chate.p@chula.ac.th

บทคัดย่อ

โปรโตคอลเว็บซ็อกเก็ต เป็นโปรโตคอลการส่งข้อมูลที่มีความเป็นไปได้ที่จะสามารถนำมาพัฒนาเป็นระบบอัปโหลดไฟล์ที่รวดเร็วและยืนยันความถูกต้องของข้อมูลได้ บทความทางวิชาการนี้นำเสนอระบบจัดการอัปโหลดไฟล์ ต้นแบบที่ใช้โปรโตคอลเว็บซ็อกเก็ตในการอัปโหลดไฟล์จากแอนดรอยด์สมาร์ทโฟนไปยังเซิร์ฟเวอร์ การทดสอบแสดงให้เห็นว่ามีความเร็วสูงกว่าเอชทีทีพีในกรณีของการอัปโหลดไฟล์ขนาดใหญ่

คำสำคัญ: การส่งไฟล์ เว็บซ็อกเก็ต แอนดรอยด์ เครื่องข่ายไร้สาย

Abstract

Websocket protocol is a network transport protocol that is possible to develop into file upload system that fast data transfer and confirm data integrity. In this paper, we propose a file upload manager prototype system that uses Websocket protocol to upload file from android smartphone to server. Experiment result shows that Websocket is faster than HTTP in case of large file upload.

Keywords: File Transfer, Websocket, Android, Wireless Network.

1. บทนำ

แนวโน้มในการใช้สมาร์ทโฟนเพื่อทำการอัปโหลดไฟล์ไปยังเซิร์ฟเวอร์นั้นมีการใช้งานที่แพร่หลายมากขึ้น ตัวอย่างโปรแกรมที่มีความสามารถดังกล่าวตัวอย่างเช่น เฟซบุ๊ก (Facebook) ดรอปบ็อกซ์ (Dropbox) เป็นต้น การอัปโหลดไฟล์ควรจะต้องใช้แบนวิธอย่างคุ้มค่า สามารถอัปโหลดได้อย่างรวดเร็ว รวมถึงตรวจสอบความถูกต้องของข้อมูลด้วยเช่นกัน

การส่งข้อมูลผ่านโปรโตคอลเอชทีทีพี (HTTP) [1] หรือโปรโตคอลเอฟทีพี (FTP) [2] เป็นวิธีการอัปโหลดไฟล์ที่ใช้ทั่วไปสำหรับอุปกรณ์สมาร์ทโฟน ในปี 2011 ได้มีการนำเสนอโปรโตคอลเว็บซ็อกเก็ต (Websocket) ซึ่งรองรับการส่งข้อมูลฐานสอง (Binary) จึงเป็นไปได้ที่จะนำโปรโตคอลนี้มาใช้ในการอัปโหลดไฟล์ได้เช่นกัน แต่เอพีไอ (API) ของเว็บซ็อกเก็ตนั้นนำมาใช้ในการอัปโหลดไฟล์ได้ยาก เพราะต้องทำการจัดการข้อมูลรับส่ง และแปลงเป็นไฟล์ด้วยตัวเอง

บทความทางวิชาการนี้จึงมีวัตถุประสงค์เพื่อนำเสนอระบบต้นแบบ (Prototype) จัดการอัปโหลดไฟล์จากแอนดรอยด์สมาร์ทโฟนไปยังเซิร์ฟเวอร์ผ่านโปรโตคอลเว็บซ็อกเก็ต ที่มีความสามารถเลือกไฟล์จากในอุปกรณ์สมาร์ทโฟนที่จะอัปโหลด จัดการรายการระหว่างอัปโหลดไฟล์ได้ ตั้งเวลาที่จะอัปโหลดไฟล์ได้ จากนั้นแสดงผลการทดสอบประสิทธิภาพเวลาที่ใช้และความเร็วในการอัปโหลด โดยอ้างอิงการตั้งค่าที่เหมาะสมที่ได้จากการทดลองที่ผ่านมา

2. ทฤษฎีที่เกี่ยวข้อง

2.1 โปรโตคอลเว็บซ็อกเก็ต

เว็บซ็อกเก็ตเป็นโปรโตคอลการสื่อสารในเครือข่ายแบบสองทางเต็มอัตรา (Full-Duplex) ไออีทีเอฟ (IETF – Internet Engineering Task Force) เป็นผู้กำหนดเป็นมาตรฐานอาร์เอฟซี 6455 เพื่อการสื่อสารแบบ โคลแอนต์-เซิร์ฟเวอร์ การทำงานคือเมื่อมีการตกลงจับมือ (Handshake) กันระหว่างอุปกรณ์โปรโตคอลเอชทีทีพีเดิม จะเปลี่ยนเป็นการส่งข้อมูลด้วยโปรโตคอลเว็บซ็อกเก็ต หลักการคล้ายกับการสื่อสารแบบซ็อกเก็ตอื่น ๆ เช่น จาวาซ็อกเก็ต (Java Socket) มีเอพีไอของภาษาต่าง ๆ เช่น จาวา (Java) จาวาสคริปต์ (Javascript)

เป็นต้น ข้อดีของเว็บซ็อกเก็ต เช่น เมื่อเปรียบเทียบกับ โพรโตคอลเอชทีทีพี พบว่าขนาดของข้อมูลส่วนหัว (Header) ในเวลาส่งแพคเกจข้อมูลของเว็บซ็อกเก็ตมีขนาดเพียง 2 ไบต์ [3] เล็กกว่าข้อมูลส่วนหัวของเอชทีทีพีมาก และเว็บซ็อกเก็ตมีค่าเวลาแฝง (Latency) ที่น้อยกว่าโพลลิง (Polling) และลองโพลลิง (Long-Polling) ในการส่งข้อมูลแบบเดียวกัน [4]

2.1 ระบบปฏิบัติการแอนดรอยด์ (Android Operating System)

แอนดรอยด์ [5] เป็นระบบปฏิบัติการของอุปกรณ์สมาร์ตโฟนและแท็บเล็ต ซึ่งเป็นที่นิยมในอุปกรณ์มือถือหลากหลายยี่ห้อหลายรุ่น เปิดช่องทางให้ผู้พัฒนาเป็นรูปแบบโอเพ่นซอร์ส (Open-source) มีเอพีไอภาษาจาวา สำหรับใช้ในการพัฒนาโปรแกรมของแอนดรอยด์

3. งานที่เกี่ยวข้อง

งานวิจัยพัฒนาระบบดูแลสุขภาพของโรงพยาบาล [4] ได้พัฒนาระบบที่มีโคลแอนต์เป็นสมาร์ตโฟนแอนดรอยด์ และมีเซิร์ฟเวอร์จัดการส่งกระจายข้อมูลรูปแบบพับลิช/ซบสไคร็บ (Publish/Subscribe) ผู้วิจัยทำการทดลองวัดผลในแง่ของเวลาเปรียบเทียบกับโพลลิง ได้ผลว่าเว็บซ็อกเก็ตตอบสนองเร็วกว่าโพลลิง ไบนารีเจเอส (BinaryJS) [6] โอเพ่นซอร์สไลบรารี (Library) ของการส่งข้อมูลสำหรับเว็บแอปพลิเคชัน (Web Application) พัฒนาต่อยอดเอพีไอพื้นฐานของเว็บซ็อกเก็ตให้สามารถส่งข้อมูลรูปแบบสตรีมมิ่งข้อมูลฐานสองได้ ซึ่งพัฒนาด้วยโหนดเจเอส (Node.js) ที่เป็นภาษาจาวาสคริปต์ ช่วยให้ผู้พัฒนาอื่นสามารถนำไปใช้ในการพัฒนาโปรแกรมต่อไปโดยไม่ต้องจัดการการรับส่งข้อมูลเองผ่านเอพีไอพื้นฐานของเว็บซ็อกเก็ต แต่ผู้พัฒนายังไม่ได้พัฒนาให้รองรับแพลตฟอร์ม (Platform) สมาร์ตโฟน งานวิจัย [7] นำเสนอวิธีการส่งข้อมูลระหว่างโคลแอนต์-เซิร์ฟเวอร์ โดยใช้จาวาซ็อกเก็ตในการส่งข้อมูลจากโคลแอนต์ที่เป็นแอนดรอยด์สมาร์ตโฟน แต่งานวิจัยนี้ไม่ได้วัดประสิทธิภาพในการส่งข้อมูล และกล่าวในบทสรุปว่า ยังไม่ได้คำนึงถึงความถูกต้องของข้อมูล โปรแกรมจัดการดาวน์โหลด แฟลชเก็ต (Download Manager – Flashget) [8] เป็นโปรแกรมจัดการดาวน์โหลดไฟล์ผ่านเครือข่าย จากการดาวน์โหลดไฟล์ผ่านเบราว์เซอร์ (Browser) ไม่สามารถดาวน์โหลด

ใหม่ต่อจากเดิมได้เมื่อเกิดข้อผิดพลาด แฟลชเก็ตสามารถดาวน์โหลดไฟล์ได้หลายไฟล์พร้อมกัน สามารถดาวน์โหลดต่อจากข้อมูลเดิมได้ สามารถสร้างหลายการเชื่อมต่อต่อหนึ่งการดาวน์โหลดเพื่อเพิ่มความเร็วได้ และสามารถตั้งเวลาในการดาวน์โหลดได้

4. ขั้นตอนวิธีการดำเนินงาน

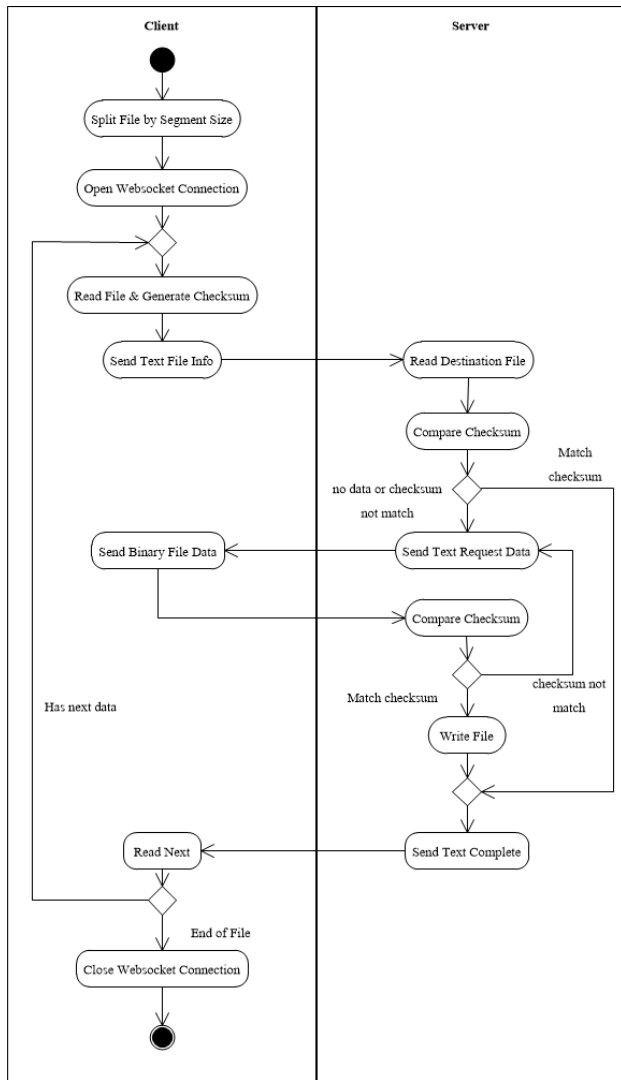
4.1 ออกแบบแนวคิด

การคำนึงถึงความถูกต้องของข้อมูลที่ได้รับ ยืนยันว่าข้อมูลของไฟล์ที่ได้รับนั้นถูกต้อง ทำโดยการตรวจสอบค่าผลรวมตรวจสอบ (Checksum) โดยเวลาก่อนส่งข้อมูลฐานสอง ระบบฝั่งโคลแอนต์จะต้องอ่านไฟล์เป็นข้อมูลค่าฐานสอง และนำค่านั้นเข้าแฮชฟังก์ชัน (Hash Function) เพื่อให้ได้ค่าผลรวมตรวจสอบ จากนั้นจึงส่งค่าผลรวมตรวจสอบให้เซิร์ฟเวอร์เก็บไว้ พอเซิร์ฟเวอร์ได้รับข้อมูลฐานสอง นำข้อมูลเข้าแฮชฟังก์ชันด้วยวิธีการเข้ารหัสแบบเดียวกันกับที่ฝั่งโคลแอนต์ใช้ ผลลัพธ์ที่ได้ นำมาเปรียบเทียบกับค่าผลรวมตรวจสอบที่ได้รับก่อนหน้า ถ้าค่าผลรวมตรวจสอบตรงกัน ก็สามารถยืนยันได้ว่า ข้อมูลที่ได้รับนั้นถูกต้องครบถ้วน

การใช้งานแบบวิธีได้อ่างมีประสิทธิภาพ ทำโดยระบบจะต้องสามารถส่งข้อมูลต่อได้ โดยที่ไม่ต้องเริ่มต้นส่งข้อมูลใหม่ตั้งแต่ต้น กรณีที่เกิดการส่งข้อมูลผิดพลาด เช่นขาดการเชื่อมต่อกันระหว่างทำการส่งข้อมูล เราจะใช้หลักการของค่าผลรวมตรวจสอบเข้ามาช่วย โดยในกรณีที่เซิร์ฟเวอร์มีข้อมูลไฟล์อยู่ส่วนหนึ่งแล้ว เมื่อเซิร์ฟเวอร์ได้รับข้อความที่มีค่าผลรวมตรวจสอบ เซิร์ฟเวอร์ก็จะอ่านไฟล์ที่มีอยู่ในตำแหน่งที่ตรงกันและขนาดเท่ากับข้อมูลที่จะได้รับ นำข้อมูลที่อ่านได้มาเข้าแฮชฟังก์ชันเพื่อหาค่าผลรวมตรวจสอบ หรือในกรณีที่มีการจัดเก็บรายการข้อมูลที่เซิร์ฟเวอร์ได้รับลงฐานข้อมูล (รวมถึงค่าผลรวมตรวจสอบ) ก็สามารถใช่วิธีอ่านค่าผลรวมตรวจสอบจากฐานข้อมูลได้เช่นกัน ตรวจสอบว่าค่าผลรวมตรวจสอบบนเซิร์ฟเวอร์ตรงกันกับค่าผลรวมตรวจสอบที่ได้รับหรือไม่ ถ้าค่าตรงกันก็ยืนยันได้ว่าเซิร์ฟเวอร์มีข้อมูลส่วนที่กำลังจะส่งมา ก็ส่งข้อมูลตอบกลับให้โคลแอนต์ข้ามการส่งข้อมูลฐานสองในส่วนนี้ไป ไม่ต้องสิ้นเปลืองแบนวิธในการส่งข้อมูลเดิมขึ้นมา

4.2 ออกแบบขั้นตอนในการอัปโหลดไฟล์

ขั้นตอนการส่งข้อมูลแสดงดังภาพที่ 1



ภาพที่ 1 : แผนภาพกิจกรรม (Activity Diagram) ขั้นตอนการติดต่อสื่อสารเพื่ออัปโหลดไฟล์

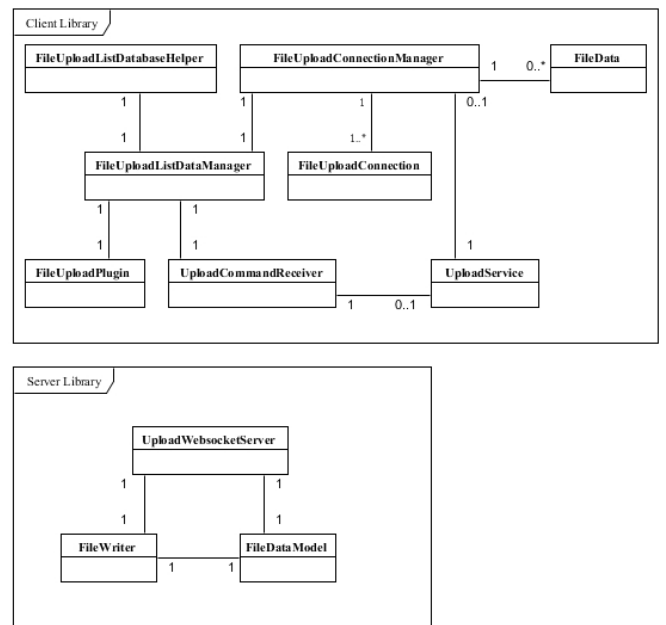
ขั้นตอนการสื่อสารระหว่างไคลเอนต์กับเซิร์ฟเวอร์ เริ่มจากไคลเอนต์แบ่งไฟล์ สร้างการเชื่อมต่อเว็บซ็อกเก็ต เริ่มส่งข้อมูลโดยอ่านไฟล์ทีละส่วน ส่งข้อมูลประเภทข้อความบอกรายละเอียดของไฟล์ ตำแหน่งของข้อมูลฐานสองที่จะเขียนลงไฟล์ และค่าผลรวมตรวจสอบก่อน เมื่อเซิร์ฟเวอร์ได้รับ ก็ตรวจสอบค่าผลรวมตรวจสอบกับไฟล์ที่มีอยู่ ถ้าข้อมูลตรงกันก็ตอบกลับให้ข้ามการส่งข้อมูลฐานสองในส่วนนี้ไปได้ แต่ถ้าค่าผลรวมตรวจสอบไม่ตรงหรือไม่มีข้อมูล เซิร์ฟเวอร์จึงตอบกลับให้ทำการส่งข้อมูลตามที่ตั้งไว้ เมื่อได้รับข้อมูลฐานสองจึงนำ

ข้อมูลเข้าแซงฟังก์ชัน และเปรียบเทียบค่าผลรวมตรวจสอบที่ได้รับก่อนหน้า ถ้าตรงกันก็นำข้อมูลเขียนลงไฟล์และตอบกลับให้ส่งข้อมูลส่วนถัดไป ถ้าไม่ตรงกันก็ตอบกลับให้ส่งข้อมูลส่วนนี้ใหม่ วนซ้ำจนได้ข้อมูลครบ จึงปิดการเชื่อมต่อ

ส่งข้อมูลหลายส่วนพร้อมกัน ด้วยการเชื่อมต่อเว็บซ็อกเก็ตมากกว่า 1 ช่อง ซึ่งคาดว่าจะช่วยลดเวลาการรวมในการอัปโหลดไฟล์ได้ ทั้งฝั่งไคลเอนต์และเซิร์ฟเวอร์จะต้องจัดการรับส่งข้อมูลหลายการเชื่อมต่อได้ เวลาอ่านหรือเขียนไฟล์ ต้องจัดการไม่ให้เกิดการล็อกไฟล์ หรือเขียนข้อมูลผิดตำแหน่ง

4.3 ออกแบบระบบ

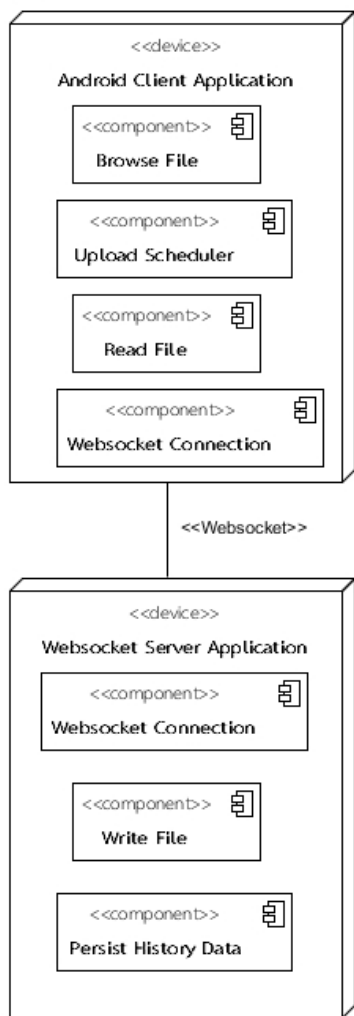
แผนภาพคลาสแสดงดังภาพที่ 2



ภาพที่ 2 : แผนภาพคลาส (Class Diagram) ไลบรารีของระบบ

ฝั่งไคลเอนต์ ส่วนต่อประสานจะเรียกใช้งานไลบรารีผ่าน FileUploadPugin ไปยัง UploadCommandReceiver และตั้ง UploadService ให้สร้าง FileUploadConnectionManager ทำการแบ่งอ่านไฟล์ ทำแซงฟังก์ชันข้อมูล เก็บใน FileData สร้าง FileUploadConnection ทำการเชื่อมต่อเว็บซ็อกเก็ตและรับส่งข้อมูล FileUploadListDataManager บันทึกข้อมูลลงฐานข้อมูล โดยเรียกใช้คำสั่งภาษาเอสคิวแอล (SQL) ผ่าน FileUploadListDatabaseHelper ฝั่งเซิร์ฟเวอร์ จะมี UploadWebsocketServer คอยรับการเชื่อมต่อ FileWriter จัดการเขียนไฟล์ และ FileDataModel บันทึกลงฐานข้อมูล

ภาพรวมระบบแสดงดังภาพที่ 3



ภาพที่ 3 : แผนภาพการติดตั้ง (Deployment Diagram) ภาพรวมของระบบ

ส่วนของไคลแอนต์ เป็น โปรแกรมประยุกต์แอนดรอยด์ ที่มีความสามารถในการเลือกไฟล์ภายในอุปกรณ์ มีส่วนจัดการรายการไฟล์ที่จะอัปโหลด และกำหนดตารางเวลาของการอัปโหลดได้ เมื่ออัปโหลด จะอ่านไฟล์ที่เลือกอัปโหลด จัดการอ่านรายละเอียดข้อมูลของไฟล์ รวมถึงจัดการค่าผลรวมตรวจสอบของการอ่านข้อมูลไฟล์ และสร้างการเชื่อมต่อเว็บซ็อกเก็ต คอยควบคุมการเชื่อมต่อการรับส่งข้อมูล

ส่วนของเซิร์ฟเวอร์ เป็น โปรแกรมเว็บซ็อกเก็ตเซิร์ฟเวอร์คอยรับการเชื่อมต่อเว็บซ็อกเก็ต เมื่อได้รับก็จะคอยควบคุมการเชื่อมต่อ การรับส่งข้อมูล คอยทำการเขียนไฟล์จากข้อมูลที่ฝั่งไคลแอนต์ส่งมา และจัดเก็บประวัติการรับส่งข้อมูลรายละเอียดของข้อมูลไฟล์ที่ได้รับลงฐานข้อมูลของเซิร์ฟเวอร์

เพื่อให้เซิร์ฟเวอร์นั้นสามารถตรวจสอบข้อมูลย้อนหลังได้ ซึ่งจำเป็นกับการส่งข้อมูลต่อจากไฟล์ในกรณีที่เกิดข้อผิดพลาด

4.4 พัฒนาระบบ

รูปแบบการส่งข้อความบอกรายละเอียดไฟล์ จะส่งในรูปแบบของเจสัน (JSON) บอกชื่อไฟล์(name) ขนาดไฟล์ (size) ขนาดชิ้นข้อมูล (chunksize) ตำแหน่งไบต์เริ่มต้นของข้อมูล (start) ตำแหน่งไบต์สิ้นสุดของข้อมูล (stop) ค่าผลรวมตรวจสอบ (checksum) และหมายเลขอุปกรณ์ (deviceId) ตัวอย่างเช่น {"stop":12671,"start":0,"name":"t-table.pdf","chunksize":12671,"checksum":"dca5b5217994a22254f5881d8dd10246","deviceId":"42a45f9d226d9dc8","size":12671} เป็นต้น

พัฒนาไคลแอนต์ด้วยภาษาจาวาและโฟนแกปเอพีไอ (PhoneGap API) ไลบรารีเว็บซ็อกเก็ตโอโตบาร์น (Autobahn) ฐานข้อมูลเฮสคิวไลท์ (SQLite) และพัฒนาเซิร์ฟเวอร์ด้วยโหนดเจส ไลบรารีเว็บซ็อกเก็ตคัมบลิเวส (ws) ฐานข้อมูลมอโกดีบี (MongoDB)

4.5 ออกแบบการทดลอง

ทำการทดลองโดยติดตั้งระบบต้นแบบฝั่งไคลแอนต์ลงแอนดรอยด์สมาร์ตโฟน และติดตั้งระบบต้นแบบฝั่งเซิร์ฟเวอร์ลงคอมพิวเตอร์เครื่อง ใช้ไฟล์ขนาดแตกต่างกัน และทดลองภายใต้สภาพเครือข่ายที่ต่างกัน ได้แก่ เครือข่ายไร้สายแลน (LAN) เดียวกัน เครือข่าย 3 จี และเครือข่าย 2 จี วัดค่าเวลาที่ใช้ในการอัปโหลดไฟล์ หากเฉลี่ย ค่ามากที่สุด ค่าน้อยสุด และค่าเบี่ยงเบนมาตรฐาน รวมถึงแปรผลเป็นค่าความเร็วอัปโหลด กำหนดค่าขนาดชิ้นของข้อมูลที่ตัดแบ่งและจำนวนการเชื่อมต่อจากค่าที่เหมาะสมที่สุดตามแต่ละเครือข่าย ซึ่งได้มาจากการทดลองอัปโหลดไฟล์ขนาดเท่าเดิม ปรับเปลี่ยนค่าขนาดชิ้นข้อมูลและจำนวนการเชื่อมต่อ เพื่อหาว่าควรจะต้องตั้งค่าอย่างไรจึงจะเหมาะสมที่สุด (ไม่ได้หมายถึงว่าเร็วที่สุดเสมอ แต่มีความเร็วสูงลำดับต้นโดยไม่ใช้จำนวนการเชื่อมต่อเกินความจำเป็นด้วยเช่นกัน) พบว่าเครือข่ายแลนไร้สาย ใช้ขนาดชิ้นข้อมูล 2 เมกะไบต์ 4 การเชื่อมต่อ ได้ความเร็วเฉลี่ย 2.15 เมกะไบต์ต่อวินาที เครือข่าย 3 จี ใช้ขนาดชิ้นข้อมูล 2 เมกะไบต์ 1 การเชื่อมต่อ ได้ความเร็วเฉลี่ย 370 กิโลไบต์ต่อวินาที และเครือข่าย 2 จี ใช้ขนาดชิ้นข้อมูล 64 กิโลไบต์ 1 การเชื่อมต่อ ได้ความเร็วเฉลี่ย 7 กิโลไบต์ต่อวินาที

กรณีของการทดลองในเครือข่ายแลนไร้สาย จะนำค่าเวลาที่นำมาเปรียบเทียบกับค่าเวลาที่ใช้ของการส่งผ่านโปรโตคอลเอชทีทีพี ข้อมูลจากระบบต้นแบบที่ใช้ไลบรารี FileTransfer ของโฟนแอปเอพีไอ นำมาแปรผลเปรียบเทียบค่าความเร็วเฉลี่ย และพิสูจน์ว่ามีความแตกต่างอย่างมีนัยยะสำคัญหรือไม่ โดยใช้การทดสอบทางสถิติ การทดสอบทีของสตีวเคนท์ (Student's t-test) ส่วนข้อมูลเวลาของเครือข่าย 3 จี และของเครือข่าย 2 จี นั้นขอยกเว้น เนื่องจากข้อจำกัดทางด้านทรัพยากร

5. ผลการดำเนินงาน

5.1 ผลการทดสอบด้านเวลาและความเร็วในการอัปโหลด

ผลการทดสอบเวลาที่ใช้วัดเป็นหน่วยมิลลิวินาที (Millisecond - ms) และความเร็ววัดเป็นหน่วยเมกะไบต์ต่อวินาที (Megabyte per second – MB/s) ในการอัปโหลดไฟล์ได้ค่าเฉลี่ย (Average) ค่ามากที่สุด (Max) ค่าน้อยสุด (Min) ค่าเบี่ยงเบนมาตรฐาน (Standard Deviation - SD) ของเครือข่ายแลนไร้สาย เครือข่าย 3 จี เครือข่าย 2 จี แสดงดังตารางที่ 1 ตารางที่ 2 และตารางที่ 3 ตามลำดับ

ตารางที่ 1 : ผลทดสอบเวลาที่ใช้และความเร็วบนเครือข่ายแลนไร้สาย

ขนาดไฟล์		2.5 MB	10 MB	100 MB	500 MB
เวลา (ms)	Average	1698.4	5460	43280	213367
	Max	1867	6227	47879	224623
	Min	1500	4975	41517	209313
	SD	131.81	473.58	1996.67	4321.98
ความเร็ว (MB/s)	Average	1.472	1.831	2.311	2.343
	Max	1.667	2.010	2.409	2.389
	Min	1.339	1.606	2.089	2.226

ตารางที่ 2 : ผลทดสอบเวลาที่ใช้และความเร็วบนเครือข่าย 3 จี

ขนาดไฟล์		2.5 MB	10 MB	50 MB	100 MB
เวลา (ms)	Average	26279	42688	282098	407827
	Max	29475	44250	302098	417827
	Min	23485	42276	262098	387827
	SD	2165.0	873.4	14142.1	11726.9
ความเร็ว (MB/s)	Average	0.095	0.234	0.177	0.245
	Max	0.106	0.237	0.191	0.258
	Min	0.085	0.226	0.166	0.239

ตารางที่ 3 : ผลทดสอบเวลาที่ใช้และความเร็วบนเครือข่าย 2 จี

ขนาดไฟล์		2.5 MB	10 MB	20 MB
เวลา (ms)	Average	430368	1570916	3288739
	Max	446368	2071076	3688729
	Min	406668	1071110	3055729
	SD	14621.73	412325.76	252674.69
ความเร็ว (MB/s)	Average	0.006	0.006	0.006
	Max	0.006	0.009	0.007
	Min	0.006	0.005	0.005

จากผลการทดสอบของเว็บซ็อกเก็ตบนเครือข่ายแลนไร้สาย นำมาเปรียบเทียบกับผลการทดสอบของเอชทีทีพีบนเครือข่ายแลนไร้สาย เพื่อหาคำตอบของสมมุติฐานว่ามีความแตกต่างระหว่างเว็บซ็อกเก็ตกับเอชทีทีพีหรือไม่ ตั้งสมมุติฐานว่าง (null hypothesis – H_0) และสมมุติฐานทางเลือก (alternative hypothesis – H_a) ไว้ดังนี้

H_0 : ความเร็วในการอัปโหลดไฟล์ไม่แตกต่างกันระหว่างเว็บซ็อกเก็ตกับเอชทีทีพี

H_a : ความเร็วในการอัปโหลดไฟล์แตกต่างกันระหว่างเว็บซ็อกเก็ตกับเอชทีทีพี

ใช้การทดสอบทีของสตีวเคนท์ หาค่าความน่าจะเป็น (probability P-value) แบบแจกแจงสองด้าน (2 tails) โดยกำหนดค่าระดับนัยสำคัญ (level of significant - α) ไว้ที่ 0.1 ผลค่า P-value แสดงดังตารางที่ 4

ตารางที่ 4 : ผลความเร็วและค่า P-value เปรียบเทียบระหว่างเว็บซ็อกเก็ตกับเอชทีทีพีบนเครือข่ายแลนไร้สาย

ขนาดไฟล์		2.5 MB	10 MB	100 MB	500 MB
ความเร็ว (MB/s)	Websocket	1.472	1.831	2.311	2.343
	HTTP	2.074	2.300	2.503	2.147
t-test P-value		0.0000104	0.0000002	0.0000996	0.0087030

ผลที่ได้พบว่าในกรณีเครือข่ายแลนไร้สาย ความเร็วของเว็บซ็อกเก็ตจะช้ากว่าเอชทีทีพีในกรณีของไฟล์ขนาดเล็ก แต่จะเริ่มมีความใกล้เคียงจนกระทั่งเร็วกว่าในกรณีที่ไฟล์ขนาดใหญ่ ตั้งแต่ 100 เมกะไบต์ขึ้นไป ค่า P-value ทุกค่าต่ำกว่าค่าระดับนัยสำคัญทุกผล จึงสรุปได้ว่า ปฏิเสธ H_0 ซึ่งก็คือความเร็วในการอัปโหลดไฟล์จะแตกต่างกันระหว่างเว็บซ็อกเก็ตกับเอชทีทีพีอย่างมีนัยยะสำคัญ

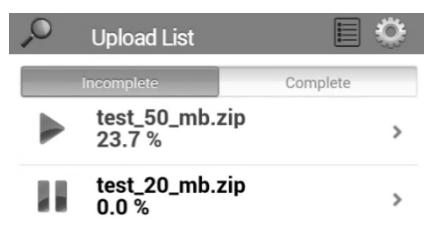
5.2 ตัวอย่างหน้าจอระบบฝั่งไคลแอนต์

หน้าจอหลักของโปรแกรมประยุกต์แอนดรอยด์ ประกอบด้วย หน้าจอเลือกไฟล์ที่เก็บภายในสมาร์ตโฟน สามารถกดที่รายการโฟลเดอร์เพื่อเปิดโฟลเดอร์ย่อยที่เลือกได้ สามารถกดที่รายการไฟล์ตรงบริเวณชื่อไฟล์เพื่อทำการอัปโหลดไฟล์นั้น ระบบจะถามว่าจะส่งไฟล์นั้นทันที เก็บในคิวรออัปโหลดทีหลัง หรือตั้งเป็นคิวรออัปโหลดตามเวลาที่ตั้งไว้ได้ สามารถเลือกหลายไฟล์พร้อมกันด้วยการกดที่รูปไฟล์เพื่อเลือก แล้วกดปุ่ม upload select เพื่อส่งเข้าคิวอัปโหลดพร้อมกันได้เช่นกัน หน้าจอแสดงรายการอัปโหลด มีส่วนแสดงความคืบหน้าของการอัปโหลด ผู้ใช้สามารถเลือกทำการส่งไฟล์ต่อ หรือหยุดการส่งไฟล์ชั่วคราวสำหรับรายการที่กำลังอัปโหลดอยู่ได้

หน้าจอของการเลือกไฟล์ หน้าจอแสดงรายการอัปโหลด แสดงดังภาพที่ 4 และภาพที่ 5 ตามลำดับ



ภาพที่ 4 : หน้าจอเลือกไฟล์



ภาพที่ 5 : หน้าจอรายการอัปโหลด

6. สรุป

จากเดิมที่เอพีไอพื้นฐานของเว็บซ็อกเก็ต ไม่รองรับกับการอัปโหลดไฟล์โดยตรง จึงต้องพัฒนาระบบที่ช่วยให้สามารถอัปโหลดไฟล์ผ่านโปรโตคอลเว็บซ็อกเก็ตได้อย่างถูกต้องและมีประสิทธิภาพ ระบบจัดการอัปโหลดสามารถเพิ่มจำนวนการ

เชื่อมต่อเพื่อให้สามารถส่งข้อมูลได้รวดเร็วยิ่งขึ้น และแก้ปัญหากรณีที่เกิดข้อผิดพลาด ผู้ใช้ไม่ต้องส่งข้อมูลใหม่ตั้งแต่ต้น ทำการทดสอบในเครือข่ายไร้สายแลนเดียวกัน 2 จี และ 3 จี เพื่อหาค่าเวลาที่ใช้และความเร็วของการอัปโหลดไฟล์ และกรณีของเครือข่ายแลนไร้สาย ได้เปรียบเทียบความเร็วกับเซททีพีและวัดผลทางสถิติ ซึ่งสรุปได้ว่า ความเร็วในการอัปโหลดไฟล์แตกต่างกันระหว่างเว็บซ็อกเก็ตกับเซททีพี โดยเว็บซ็อกเก็ตมีความเร็วสูงกว่าในกรณีไฟล์ขนาดใหญ่ งานในอนาคตคือวิเคราะห์หาทางปรับปรุงประสิทธิภาพของไลบรารี ให้อย่างน้อยมีประสิทธิภาพทัดเทียมกับเซททีพีในทุกกรณี

เอกสารอ้างอิง

- [1] A. S. Inc. (2013, July 6). *PhoneGap API Documentation File API FileTransfer Object*. Available: http://docs.phonegap.com/en/3.0.0/cordova_file_file.md.html#FileTransfer
- [2] C. E. Palazzi and M. Ferrarese, "FTP4Android: A local/remote file manager for Google Android platform," in *Consumer Communications and Networking Conference (CCNC)*, 2011 *IEEE*, 2011, pp. 373-377.
- [3] K. Corporation. (2013, July 6). *HTML5 Web Sockets: A Quantum Leap in Scalability for the Web*. Available: <http://www.websocket.org/quantum.html>
- [4] R. Lomotey, R. Kazi, and R. Deters, "Near real-time medical data dissemination in m-Health," presented at the Proceedings of the International Conference on Management of Emergent Digital EcoSystems, Addis Ababa, Ethiopia, 2012.
- [5] G. Inc. (2013, 9 March). *Android, the world's most popular mobile platform*. Available: <http://developer.android.com/about/index.html>
- [6] E. Zhang. (2013, June 11). *BinaryJS - Realtime binary streaming for the web using websockets*. Available: <http://binaryjs.com>
- [7] J. Yan Ling, Y. Zheng Xiao, D. Cao, and S. Zheng, "Development and implementation of Eclipse-based file transfer for Android Smartphone," in *Computer Science & Education (ICCSE)*, 2012 *7th International Conference on*, 2012, pp. 568-571.
- [8] T. M. C. Limited. (2013, August 20). *Best Download Manager - Flashget*. Available: <http://www.flashget.com>

การประเมินประสิทธิภาพในการรองรับมัลติมีเดียของเครือข่าย 3G: กรณีศึกษา การวัดคุณภาพเสียงจากสไกป์และเฟสบุ๊คในบริเวณกรุงเทพฯชั้นใน

Performance Evaluation of 3G Networks for Multimedia: A Case Study of Voice Quality Measurement from Skype and Facebook in the Inner City of Bangkok.

วรวิษ เอกาพันธ์ (Worawat Akapan)¹ ตูลย์ ไตรยสรรรค์ (Tuul Triyason)²

เทอดพงษ์ แดงสี (Therdpong Daengsi)³ และ พงษ์พิสิฐ วุฒิดิษฐโชติ (Pongpisit Wuttidittachotti)⁴

^{1,4}คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

²คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี

³เอ็นเตอร์ไพรส์เซอร์วิสเชส บริษัท แจ็คส์คอม จำกัด

¹dekip_555@hotmail.com, ²51501701@st.sit.kmutt.ac.th,

³therdpong1@yahoo.com, ⁴pongpisitw@kmutnb.ac.th

บทคัดย่อ

บทความวิจัยนี้นำเสนอเกี่ยวกับการวัดคุณภาพของเครือข่าย 3G ภายในพื้นที่ชั้นในกรุงเทพมหานคร โดยประเมินค่าคุณภาพเสียงที่ได้จากการทดสอบด้วยสไกป์ ซึ่งเป็นบริการวีโอไอพีแอปพลิเคชันและเฟสบุ๊ค ซึ่งเป็นเครือข่ายสังคมออนไลน์ที่มีฟังก์ชันในการให้บริการวีโอไอพี โดยทำการทดสอบและเก็บข้อมูลจาก 5 เครือข่ายหลักที่ให้บริการ 3G ในบริเวณ 7 มหาวิทยาลัยในเขตกรุงเทพฯชั้นใน แล้วนำไปวัดด้วยเครื่องวัด PESQ เพื่อหาค่า MOS และ ทำการวิเคราะห์ด้วย ANOVA และ T-test จากผลการศึกษาพบว่า สามารถจัดกลุ่มผู้ให้บริการออกเป็น 3 กลุ่ม คือ กลุ่มคุณภาพดี คือ TOT และ TRUE (วัดค่า MOS ได้ $\approx 3.8-3.9$) กลุ่มคุณภาพปานกลาง คือ CAT (วัดค่า MOS ได้ ≈ 3.7) และกลุ่มพอใช้ คือ DTAC และ AIS (วัดค่า MOS ได้ $\approx 3.1-3.6$) นอกจากนี้ยังพบว่า กลุ่มสุดท้ายครอบคลุมพื้นที่ให้บริการน้อยกว่ากลุ่มที่ 1 และกลุ่มที่ 2

คำสำคัญ: เครือข่าย 3G พีอีเอสคิว เอ็มไอเอส วีโอไอพี

Abstract

This paper presents the performance evaluation of 3G networks for multimedia within the inner city of Bangkok using Skype and Facebook. This study used evaluation of voice quality provided by Skype, a VoIP application, and Facebook, a service website that has VoIP function. The tests and data have been conducted and gathered

respectively for 5 major 3G networks, serving by 5 operators, within 7 universities in the inner city of Bangkok. Then the data were measured by using Perceptual Evaluation of Speech Quality (PESQ) to find Mean Opinion Score (MOS) before analyzing with ANOVA and T-test, which are statistical tools. From this study, it has been found that 3G networks can be classified into 3 groups, consisting of the good quality group, TOT and TRUE (MOS $\approx 3.8-3.9$), the fair quality group, CAT only (MOS ≈ 3.7), and the poor quality group, DTAC and AIS (MOS $\approx 3.1-3.6$). Besides, it has been found that the last group of 3G networks cover the test areas less than the first and the second group.

Keywords: 3G Network, PESQ, MOS, VoIP

1. บทนำ

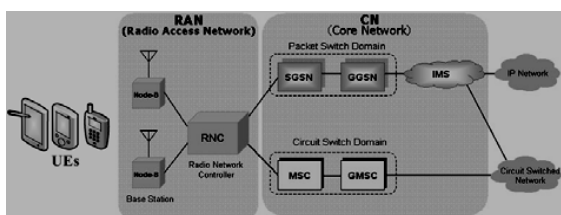
VoIP (Voice over Internet Protocol) เป็นรูปแบบการสื่อสารข้อมูลผ่านระบบอินเทอร์เน็ตที่ได้รับความนิยมเป็นอย่างมาก สามารถใช้งานร่วมกันได้ทั้งเครื่องคอมพิวเตอร์กับเครื่องคอมพิวเตอร์ เครื่องคอมพิวเตอร์กับโทรศัพท์ และ โทรศัพท์กับโทรศัพท์ ซึ่งสามารถทำงานร่วมกันได้โดยมีค่าใช้จ่ายที่มีความประหยัดกว่าโทรศัพท์พื้นฐาน แต่อย่างไรก็ตามก็ยังมีข้อจำกัดในการใช้คือ มีปัญหาเรื่องคุณภาพเสียงเมื่อเกิดการประวิงเวลา (Delay) หรือการสูญเสียข้อมูล (Loss) ในช่วงเวลาที่มีผู้ใช้งานจำนวนมาก ซึ่งเกิดขึ้นได้เสมอบนเครือข่ายอินเทอร์เน็ต [1-2]

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาประสิทธิภาพในการรองรับมัลติมีเดียของเครือข่าย 3G เพราะการประเมินจากความเร็วในการรับ-ส่งข้อมูล ที่ได้จากการทดสอบผ่าน Speed Test เพียงอย่างเดียวไม่เพียงพอ ดังนั้นจึงนำวีโอไอพีมาประเมินประสิทธิภาพโดยทำการศึกษาเปรียบเทียบคุณภาพเสียง ทั้งนี้ผู้วิจัยได้เลือกแอปพลิเคชันสไกป์และบริการวีโอไอพีบนเว็บไซต์ของเฟสบุ๊กเพื่อใช้ในการทดสอบ โดยนำผลการทดสอบที่ได้มาทำการวิเคราะห์คุณภาพเสียงด้วยเครื่องวัดคุณภาพเสียงพีอีเอสคิว (Perceptual Evaluation of Speech Quality: PESQ) บนเครือข่าย 3G ผ่านผู้ให้บริการที่แตกต่างกันในการทดสอบจะทดสอบด้วยเสียงภาษาอังกฤษแบบบริติช 8 ไฟล์ และแบบอเมริกัน 8 ไฟล์ [3] เมื่อได้ผลจากการทดสอบนำมาวิเคราะห์เปรียบเทียบเพื่อให้ทราบว่าแต่ละเครือข่ายมีประสิทธิภาพหรือคุณภาพอยู่ในระดับใด

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 เครือข่าย 3G

เครือข่าย 3G เป็นเครือข่ายโทรศัพท์เคลื่อนที่รุ่นที่ 3 (International Mobile Telecommunications 2000: IMT-2000 และ the Universal Mobile Telecommunications System: UMTS) ที่ได้ทำการเปิดตัวในปี 2001 ซึ่งกำหนดร่วมกันโดย International Telecommunication Union และ (Third Generation Partnership Project: 3GPP) [4-7] เครือข่าย 3G สามารถรองรับการให้บริการหลายอย่างรวมถึงบริการอินเทอร์เน็ตบนโทรศัพท์มือถือ ด้วยการรับส่งข้อมูลที่ ≥ 200 kbps [8] รูปที่ 1 จะเห็นได้ว่าเครือข่าย 3G จะมีโครงสร้างแบบต้นไม้ซึ่งประกอบด้วย อุปกรณ์ของผู้ใช้ สถานีฐาน ส่วนควบคุมสัญญาณ จีอาร์เอสสนับสนุนโหนด และเกตเวย์จีพีอาร์เอส สนับสนุนโหนด จุดเชื่อมต่อโทรศัพท์ เกตเวย์จุดเชื่อมต่อโทรศัพท์ และ ไอพีมัลติมีเดีย ซับสเตชัน (IP Multimedia Substation: IMS) ซึ่งเป็นส่วนสำคัญในระบบ 3G ที่ทำหน้าที่รองรับมัลติมีเดีย [9-10]



รูปที่ 1: องค์ประกอบของเครือข่าย 3G

2.2 มัลติมีเดีย: สไกป์และเฟสบุ๊ก

สไกป์และเฟสบุ๊ก เป็นมัลติมีเดียต่างชนิดกัน โดยสไกป์จะเป็นวีโอไอพีแอปพลิเคชัน ส่วนเฟสบุ๊กเป็นเว็บไซต์บริการที่มีฟังก์ชัน วีโอไอพี

สไกป์ (Skype) เป็น โปรแกรมที่ใช้คุยโทรศัพท์ที่สามารถคุยแบบวีดีโอหรือส่งข้อความผ่านอินเทอร์เน็ต มีคุณภาพเสียงที่ชัดเจน และไม่มีค่าใช้จ่าย ถ้าโทรผ่านคอมพิวเตอร์ ผู้ใช้สามารถโทรหาเบอร์โทรศัพท์และส่งข้อความได้ [12]

เฟสบุ๊ก (Facebook) เป็นเครือข่ายสังคมที่นิยมใช้กันอย่างแพร่หลาย เฟสบุ๊กยังมีแอปพลิเคชันที่สามารถใช้วีดีโอคอลและส่งข้อความผ่านเว็บไซต์ของเฟสบุ๊กเองได้โดยไม่ต้องเสียค่าใช้จ่าย และยังมีลูกเล่นต่างๆ ที่สามารถดึงดูดกลุ่มผู้ใช้งานจนติดอันดับเว็บไซต์ยอดนิยมในปัจจุบัน [13]

2.3 การวัดคุณภาพเสียง

การวัดคุณภาพเสียงมี 2 ประเภท คือ การวัดเชิงจิตวิสัย (Subjective Measurement) และการวัดเชิงวัตถุวิสัย (Objective Measurement) ทั้งนี้ ITU-T ได้นำเสนอมาตรที่ใช้วัดคุณภาพเสียงวีโอไอพีด้วยการหาค่า Mean Opinion Score (MOS) ซึ่งทำได้โดยการทดสอบกับคนจำนวนมากพอสมควร แล้วทำการหาคะแนนคุณภาพเสียง (MOS) เป็นค่าเฉลี่ยแต่ก็มีข้อจำกัด คือเป็นวิธีที่ใช้เวลาค่อนข้างนานในการทดสอบและเก็บข้อมูล เพราะต้องใช้คนจำนวนมากในการทดสอบ ดังนั้น ITU-T จึงได้นำเสนอวิธีการวัดคุณภาพเสียงเชิงวัตถุวิสัย (Objective) ซึ่งหนึ่งในวิธีวัดคุณภาพเสียงประเภทนี้ที่เป็นที่นิยม คือ การวัดแบบพีอีเอสคิว [14] โดยปกติค่า MOS หรือค่าคุณภาพเสียงมีค่าอยู่ระหว่าง 1 ถึง 5 (5 หมายถึงคุณภาพที่ดีที่สุด และลดลงจนถึงระดับ 1 ซึ่งหมายถึงค่าที่ไม่สามารถยอมรับได้) [15]

2.4 งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวกับการเปรียบเทียบคุณภาพเสียงของวีโอไอพี มีงานวิจัยที่ศึกษาหลายงานด้วยกัน ซึ่งแต่ละงานวิจัยใช้วิธีที่แตกต่างกันโดยขึ้นอยู่กับเงื่อนไขในแต่ละงานวิจัยนั้นๆ ดังนี้

- L. Sun และ E. Ifeachor [16] ได้ทำการทดสอบคุณภาพเสียงมากกว่า 1800 ไฟล์เสียงโดยทดสอบแบบ Mobile to Mobile และ Mobile to PSTN โดยใช้เวลาสามเดือน ผลที่ได้รับจะพบปัญหาเมื่อนำมาทดสอบกับเครื่องวัดแบบพีอีเอสคิว ปัญหาที่เกิดขึ้นใหญ่จะเกิดกับตำแหน่งของ

ช่วงเวลาที่ ไม่เหมาะสมในการวัดคุณภาพเสียงแบบพีอีเอสคิว เมื่อมีช่องว่าง และเสียงแทรกจากเสียงตัวอย่าง

- K. Yochanang และคณะ [17] ได้ทำการศึกษาคุณภาพเสียงที่ได้จาก G.711 และ G.729 ด้วยเครื่องวัด พีอีเอสคิว และเสียงภาษาไทย จำนวน 4 ชุดย่อย พบว่า ค่าคุณภาพเสียงจากแต่ละชุดย่อยไม่แตกต่างกันอย่างมีนัยสำคัญ และพบว่า เสียงผู้ชายให้ค่าคุณภาพเสียงดีกว่าเสียงผู้หญิง ในขณะที่เสียงเด็กให้ค่าคุณภาพเสียงต่ำที่สุด

- B. Sat และ B. Wah [18] ได้ทำการประเมินคุณภาพเสียงวีโอไอพีจากการฟังและสนทนาโต้ตอบ โดยใช้แอปพลิเคชัน ได้แก่ Skype, Google Talk, Windows Live Messenger และ Yahoo Messenger ด้วยเครื่องวัดพีอีเอสคิว โดยใช้การวัดแบบวัดตัวสัญญาณได้เงื่อนไขการประวิงเวลา ผลการทดสอบได้ว่า Windows Live Messenger มีประสิทธิภาพมากกว่าแอปพลิเคชันอื่นๆ ในขณะที่มีการสูญเสียแพ็คเก็ตเกิดขึ้น

- Z. Jianming และ Y. Dacheng [19] ได้ทำการศึกษาตรวจสอบคุณภาพของ VoIP บนเครือข่ายอินเทอร์เน็ตไร้สาย โดยนำ Gilbert Model มาใช้ในการจำลอง ในการประเมินผล ด้วยเครื่องวัดพีอีเอสคิว ผลที่ได้แสดงให้เห็นว่าอัตราการสูญเสียแพ็คเก็ตอย่างต่อเนื่องมีผลกระทบต่อคุณภาพเสียง

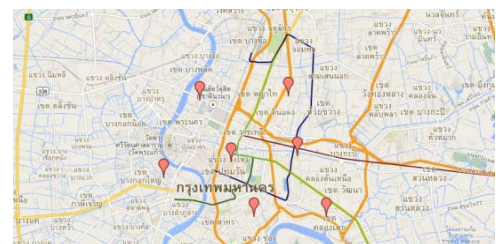
- B. Rodrigo [20] ได้ประเมินคุณภาพเสียงของ Skype และ Google Talk ในสภาพเครือข่ายที่มีการสูญเสียแพ็คเก็ต เพื่อประเมินคุณภาพของเสียงวีโอไอพีโดยใช้เครื่องวัดแบบพีอีเอสคิว และนำมาเปรียบเทียบกับอัลกอริทึมของ ITU-T ผลที่ได้ Skype มีคุณภาพเสียงที่ดีกว่าในสภาพเครือข่ายที่เกิดการสูญเสียแพ็คเก็ต แต่ก็ยังมีความคลาดเคลื่อนเกิดขึ้น

- เดือนกันยายน 2556 ส่วนงานกำกับดูแลกิจการโทรคมนาคม สำนักงานสำนักงานคณะกรรมการกิจการกระจายเสียงกิจการโทรทัศน์และกิจการโทรคมนาคมแห่งชาติ (กสทช.) [21] ได้รายงานผลการวัดคุณภาพการให้บริการโทรศัพท์เคลื่อนที่ 3G แบบเคลื่อนที่ โดยมีค่าทรูพูทเฉลี่ย (Average Data Throughput) เป็นหนึ่งในตัวชี้วัดเมื่อทดสอบด้วยการดาวน์โหลดไฟล์ข้อมูลจากเอฟทีพี

เซิร์ฟเวอร์ (FTP Server) ซึ่งไม่ใช่มัลติมีเดีย และสรุปว่า AIS ให้คุณภาพในการบริการดีที่สุด

- ช่วงต้นปี 2557 นฤมล ชุมคช และคณะ [22] ได้รายงานการศึกษาเปรียบเทียบคุณภาพเสียงที่ได้จากแอปพลิเคชันไลน์และแทงโก้ ผ่านเครือข่าย 3G และเครือข่ายไวร์เลสแลน ด้วย PESQ และเสียง 3 ภาษา (ไทย จีน และอังกฤษ) จากการศึกษาดังกล่าวพบว่า แอปพลิเคชันไลน์ให้คุณภาพเสียงที่ดีกว่าแอปพลิเคชันแทงโก้ นอกจากนี้ยังพบว่า ภายใต้อินเทอร์เน็ตและแอปพลิเคชันชนิดเดียวกัน ผลที่ได้จากทั้ง 3 ภาษาไม่มีความแตกต่างกันอย่างมีนัยสำคัญ

จากการทบทวนงานวิจัยที่เกี่ยวข้องจะเห็นได้ว่า หลังจากที่มีการเปิดใช้เครือข่าย 3G ในประเทศไทยอย่างเป็นทางการ (พฤษภาคม 2556) ยังไม่มีงานวิจัยใดที่ทำการศึกษาศักยภาพของเครือข่าย 3G ในเขตกรุงเทพมหานครอย่างจริงจัง โดยเฉพาะพื้นที่กรุงเทพมหานครชั้นใน ยกเว้นงานวิจัยของ กสทช. ที่คล้ายกับงานวิจัยนี้พอสมควร แต่แตกต่างกันในแง่ของวิธีการ เพราะงานวิจัยนี้มุ่งเน้นทดสอบประสิทธิภาพในการรองรับมัลติมีเดียของเครือข่าย 3G เป็นสำคัญ



รูปที่ 2: พิกัดสถานที่ในการทดสอบซึ่งเป็นพื้นที่กรุงเทพมหานครชั้นใน

ตารางที่ 1: ตารางแสดงสถานที่และเวลาในการเก็บข้อมูล (โดยประมาณ)

วันที่/สถานที่ทำการเก็บข้อมูล	เครือข่ายที่ใช้					
	9:00-10:30	10:30-12:00	13:00-14:30	14:30-16:00	16:00-17:30	
14/2/57 มหาวิทยาลัยราชภัฏสวนสุนันทา	TRUE	TOT	CAT	AIS	DTAC	
18/2/57 มหาวิทยาลัยศรีนครินทรวิโรฒประสมมิตร	CAT	AIS	DTAC	TRUE	TOT	
20/2/57 มหาวิทยาลัยราชภัฏบ้านสมเด็จเจ้าพระยา	DTAC	TRUE	TOT	CAT	AIS	
21/2/57 จุฬาลงกรณ์มหาวิทยาลัย	TRUE	TOT	CAT	AIS	DTAC	
24/2/57 มหาวิทยาลัยหอการค้าไทย	CAT	AIS	DTAC	TRUE	TOT	
25/2/57 มหาวิทยาลัยกรุงเทพ กว๊านน้ำไท	AIS	DTAC	TRUE	TOT	CAT	
26/2/57 มหาวิทยาลัยเทคโนโลยีราชมงคลกรุงเทพ (เทคนิครุงเทพ)	DTAC	TRUE	TOT	CAT	AIS	

ตารางที่ 2: เงื่อนไขในการทดสอบและเก็บข้อมูล

การทดสอบ	การเชื่อมต่อสัญญาณโทรศัพท์ 3G					ไฟล์เสียงทดสอบ	
	AIS	DTAC	TRUE	TOT	CAT	ภาษาอังกฤษ	จำนวนรอบ
Skype (S)	S1	S2	S3	S4	S5	16 ไฟล์	2
Facebook (FB)	FB1	FB2	FB3	FB4	FB5	16 ไฟล์	2

หมายเหตุ: S1-S5 คือ การทดสอบด้วย สไลป์ ครั้งที่ 1 – 5

FB1-FB5 คือ การทดสอบด้วย เฟสบุ๊ค ครั้งที่ 1 – 5

3. วิธีการดำเนินงาน

ในการทดสอบผู้วิจัยได้ทำการออกแบบการทดสอบ โดยเลือกบริเวณที่ใช้ในการทดสอบเป็นบริเวณพื้นที่ในมหาวิทยาลัยต่างๆ ในเขตพื้นที่ชั้นในของกรุงเทพมหานครทั้งหมด 7 แห่ง (ระหว่างวันที่ 14-26 กุมภาพันธ์ 2557) ดังแสดงในรูปที่ 2 และตารางที่ 1

จากนั้นทำการจัดเตรียมเครื่องคอมพิวเตอร์จำนวนสองเครื่องพร้อมลงแอปพลิเคชันสไกป์และเฟสบุ๊ค โดยเครื่องคอมพิวเตอร์เครื่องแรกเป็นเครื่องที่ใช้จำลองการโทรออกเพื่อส่งไฟล์เสียงต้นทางผ่านแอปพลิเคชันที่ลงไว้ และคอมพิวเตอร์เครื่องที่สองเป็นตัวจำลองผู้รับสายที่ไว้สำหรับสำหรับจับเก็บไฟล์เสียงปลายทางซึ่งได้รับจากเครื่องคอมพิวเตอร์เครื่องแรกผ่านแอปพลิเคชันเดียวกัน ทั้งนี้ไฟล์เสียงภาษาอังกฤษที่ได้มาจากเว็บไซต์ ITU-T โดยผู้วิจัยเลือกใช้เสียงภาษาอังกฤษ เพราะมีจำนวนไฟล์เสียงให้ใช้ถึง 16 ไฟล์ ขณะที่ภาษาอื่นมีเพียง 8 ไฟล์ นอกจากนี้ นฤมล ชุมคช และคณะ [22] ยังได้ระบุไว้ว่าภาษาที่แตกต่างกัน ไม่ทำให้ค่าคุณภาพเสียงที่ได้จากการทดสอบ มีความแตกต่างกันอย่างมีนัยสำคัญ

ในการทดสอบนี้จะทดสอบผ่านเครือข่ายโทรศัพท์เคลื่อนที่ 3G ที่ได้รับความนิยม 5 เครือข่าย ซึ่งประกอบด้วย AIS, DTAC, TRUE, TOT และ CAT โดยเชื่อมต่อผ่านอุปกรณ์แอร์การ์ด (Aircard) ซึ่งมีเงื่อนไขการทดสอบ ดังแสดงในแสดงตารางที่ 2 เมื่อทำการทดสอบเสร็จแล้วได้นำไฟล์เสียงต้นฉบับและไฟล์เสียงที่บันทึกได้ที่ปลายทางมาทำการวัดด้วยเครื่องวัดคุณภาพเสียงแบบพีอีเอสคิว ซึ่งเป็นมาตรฐานในการวัดคุณภาพเสียงเชิงวัตถุวิสัยแบบหนึ่งซึ่งพัฒนาโดย ITU-T และเป็นระบบเดียวกันกับที่ใช้ในการศึกษาของ นฤมล ชุมคช และคณะ [22] เพื่อหาค่าคุณภาพเสียงหรือค่า MOS จากนั้นจึงนำค่า MOS ที่ได้มาทำการวิเคราะห์และเปรียบเทียบโดยใช้ ANOVA และ T-test [23-24]

4. ผลการทดสอบและการวิเคราะห์ผล

จากตารางที่ 3 และรูปที่ 3 เป็นการแสดงค่าเฉลี่ยของค่า MOS ที่ได้จากการทดสอบไฟล์เสียงภาษาอังกฤษ ข้อมูลที่ได้รับเมื่อนำมาทดสอบผ่านเครื่องวัดคุณภาพเสียงแบบพีอีเอสคิว ซึ่งจะเห็นว่าคุณภาพเสียงที่ได้รับจากการเชื่อมต่อเครือข่าย 3G ที่ให้คุณภาพเสียงที่ดีที่สุด คือ TOT ซึ่งได้ค่า MOS = 3.93 ผ่าน

แอปพลิเคชันสไกป์ รองลงมา คือ TRUE ซึ่งได้ค่า MOS = 3.89 ผ่านฟังก์ชันวีโอไอพีเฟสบุ๊ค ในขณะที่เครือข่าย 3G ที่ให้คุณภาพต่ำที่สุด คือ AIS ค่า MOS = 3.3 ผ่านฟังก์ชันวีโอไอพีเฟสบุ๊ค นอกจากนี้ยังพบว่าเครือข่าย AIS ไม่ค่อยเสถียรเท่าที่ควร (มีค่าเบี่ยงเบนมาตรฐานหรือ SD เท่ากับ 1.05 ซึ่งสูงที่สุด) โดย AIS ยังเป็นเครือข่ายที่ครอบคลุมพื้นที่น้อยที่สุดด้วยสังเกตได้จากจำนวนข้อมูลที่ได้จากการทดสอบต่ำที่สุด (N=96 สำหรับสไกป์ และ N=64 สำหรับเฟสบุ๊ค) เนื่องจากสัญญาณไม่มีความเสถียรทำให้มีปัญหาในการเชื่อมต่อ จึงไม่สามารถเก็บข้อมูลได้ครบ (N=112) ในขณะที่เครือข่าย 3G ของ DTAC ก็ให้ผลคล้ายกับของ AIS แต่มีความเสถียรมากกว่าเล็กน้อย

จากนั้นได้ทำการวิเคราะห์ทางสถิติโดยใช้ ANOVA และ t-test เพื่อวิเคราะห์เปรียบเทียบคุณภาพของแต่ละเครือข่ายว่ามีความแตกต่างกันอย่างมีนัยสำคัญหรือไม่ โดยจะใช้นัยสำคัญที่ระดับ 0.05 หรือช่วงความเชื่อมั่น 95%

จากผลการวิเคราะห์ทางสถิติดังแสดงในตารางที่ 4 จะเห็นว่าคุณภาพที่ได้ทั้ง 5 เครือข่ายแตกต่างกันสังเกตได้จากสมมติฐาน HA: P-Value= 0.001 จากนั้นทำการเปรียบเทียบ ทีละคู่พบว่า TOT และ TRUE ให้คุณภาพเสียงที่ไม่แตกต่างกันอย่างมีนัยสำคัญ (สมมติฐาน HI: P-Value = 0.244 สำหรับสไกป์และ 0.786 สำหรับเฟสบุ๊ค) และพบว่า จากเครือข่าย 3G ของ AIS และ DTAC ไม่แตกต่างกันอย่างมีนัยสำคัญ (สมมติฐาน HB: P-Value = 0.522 สำหรับสไกป์และ 0.266 สำหรับเฟสบุ๊ค) ในขณะที่คุณภาพเสียงของเครือข่าย 3G ของ CAT ไม่แตกต่างกันอย่างมีนัยสำคัญเมื่อเทียบกับของ AIS และ DTAC ในการทดสอบด้วยสไกป์ (สมมติฐาน HE: P-Value = 0.136 และ สมมติฐาน HH: P-Value = 0.382 ตามลำดับ) แต่แตกต่างกันอย่างมีนัยสำคัญเมื่อทดสอบด้วยเฟสบุ๊ค (สมมติฐาน HA: P-Value = 0.001 และสมมติฐาน HH: P-Value = 0.005 ตามลำดับ) ในขณะที่คุณภาพเสียงของ CAT เทียบกับ TOT พบว่า ไม่แตกต่างกันอย่างมีนัยสำคัญเมื่อทดสอบด้วยเฟสบุ๊ค (สมมติฐาน HK: P-Value = 0.265) แต่มีความแตกต่างกันอย่างมีนัยสำคัญเมื่อทดสอบด้วยสไกป์ (สมมติฐาน HK: P-Value = 0.026) อย่างไรก็ตามพบว่า คุณภาพเสียงของเครือข่าย 3G ของ CAT และ TRUE ไม่แตกต่างกันอย่างมีนัยสำคัญ (สมมติฐาน HI: P-Value = 0.258 และ 0.413

เมื่อทดสอบด้วยสไกป์และเฟสบุ๊คตามลำดับ) จากผลการวิเคราะห์ข้างต้นสามารถกล่าวได้ว่า เครือข่าย 3G ของ TOT และ TRUE มีประสิทธิภาพในการรองรับมัลติมีเดียใกล้เคียงกัน รองลงมา คือ CAT ในขณะที่เครือข่าย 3G ที่มีประสิทธิภาพในการรองรับมัลติมีเดียต่ำที่สุด คือ DTAC และ AIS

จากตารางที่ 5 จะเห็นได้ว่า คุณภาพเสียงที่ได้จากสไกป์และฟังก์ชันวีโอไอพีของเฟสบุ๊ค ไม่มีความแตกต่างกันอย่างมีนัยสำคัญเมื่อทดสอบภายใต้เครือข่ายเดียวกัน (P-Value > 0.05)

อย่างไรก็ดี จากผลการศึกษานี้ไม่สอดคล้องกับผลการศึกษาของ กสทช. ที่รายงานไว้ว่า AIS เป็นเครือข่าย 3G ที่ดีที่สุด [21] ซึ่งอาจเกิดจากความแตกต่างของวิธีการทดสอบ โปรโตคอลที่ใช้ทดสอบ จุดหรือบริเวณที่ทดสอบ และตัวชี้วัด นอกจากนี้ การแข่งขันกันด้วยการพัฒนาเครือข่าย 3G ของผู้ให้บริการแต่ละราย ก็เป็นอีกหนึ่งปัจจัยที่อาจทำให้ได้ผลการศึกษาที่แตกต่างกัน

5. สรุปผลและข้อเสนอแนะ

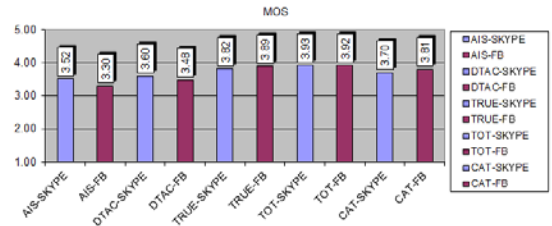
จากการศึกษานี้ ซึ่งเป็นการศึกษาเปรียบเทียบเครือข่าย 3G ของ 5 เครือข่าย สรุปได้ว่า เครือข่าย 3G ของ TOT และ TRUE มีประสิทธิภาพในการรองรับมัลติมีเดียดีที่สุดในเมื่อเทียบกับรายอื่น โดยเฉพาะอย่างยิ่งมัลติมีเดียประเภทวีโอไอพี นอกจากนี้ยังครอบคลุมทุกพื้นที่ที่ได้ทำการทดสอบ ผลจากการศึกษานี้ เป็นประโยชน์อย่างยิ่งสำหรับกลุ่มผู้ใช้งานมัลติมีเดียที่ต้องการเครือข่าย 3G ที่มีประสิทธิภาพและครอบคลุมมากที่สุด

สำหรับงานวิจัยในอนาคต เพื่อยืนยันการศึกษานี้ ควรศึกษาเพิ่มเติมโดยการเพิ่มจุด หรือสถานที่ที่ต้องทำการทดสอบและพัฒนาวิธีการในการทดสอบ

6. กิตติกรรมประกาศ

บทความนี้สำเร็จได้ด้วยความอนุเคราะห์ของมหาวิทยาลัยทั้ง 7 แห่ง ได้แก่ มหาวิทยาลัยหอการค้าไทย มหาวิทยาลัยเทคโนโลยีราชมงคลกรุงเทพ(เทคนิคกรุงเทพ) จุฬาลงกรณ์มหาวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒประสานมิตร มหาวิทยาลัยกรุงเทพ (กล้วยน้ำไท) มหาวิทยาลัยราชภัฏสวนสุนันทา และมหาวิทยาลัยราชภัฏบ้านสมเด็จเจ้าพระยาที่เอื้อเฟื้อสถานที่ในการทดลอง ส่วนคณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี ที่เอื้อเฟื้อเครื่องมือวัดพีอีเอสคิว และสุดท้ายขอขอบคุณเพื่อนๆ

ในกลุ่มงานวิจัยวีโอไอพีมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือทุกคนที่ให้ข้อมูล และคำปรึกษาเป็นอย่างดี



รูปที่ 3: ค่า MOS ที่แสดงในรูปแบบกราฟ

ตารางที่ 3: ค่า MOS ที่ได้จากการทดสอบ

VoIP App	เครือข่าย3G	MOS	SD	N
Skype	AIS	3.52	0.91	96
	DTAC	3.60	0.87	104
	TRUE	3.82	0.72	112
	TOT	3.93	0.65	112
	CAT	3.70	0.85	112
Facebook	AIS	3.30	1.05	64
	DTAC	3.48	0.85	96
	TRUE	3.89	0.71	112
	TOT	3.92	0.62	112
	CAT	3.81	0.80	112

ตารางที่ 4: ค่า P-Value จากการสมมติฐาน HA-HK

สมมติฐาน	P-Value	
	Skype	Facebook
HA ₀ คุณภาพเสียงจากแต่ละเครือข่ายเท่ากัน	< 0.001*	< 0.001*
HA ₁ คุณภาพเสียงจากแต่ละเครือข่ายไม่เท่ากัน		
HB ₀ MOS _{AIS} = MOS _{DTAC}	0.522	0.266
HB ₁ MOS _{AIS} ≠ MOS _{DTAC}		
HC ₀ MOS _{AIS} = MOS _{TRUE}	0.009*	0.000*
HC ₁ MOS _{AIS} ≠ MOS _{TRUE}		
HD ₀ MOS _{AIS} = MOS _{TOT}	0.000*	0.000*
HD ₁ MOS _{AIS} ≠ MOS _{TOT}		
HE ₀ MOS _{AIS} = MOS _{CAT}	0.136*	0.001*
HE ₁ MOS _{AIS} ≠ MOS _{CAT}		
HF ₀ MOS _{DTAC} = MOS _{TRUE}	0.044*	0.000*
HF ₁ MOS _{DTAC} ≠ MOS _{TRUE}		
HG ₀ MOS _{DTAC} = MOS _{TOT}	0.002*	0.000*
HG ₁ MOS _{DTAC} ≠ MOS _{TOT}		
HH ₀ MOS _{DTAC} = MOS _{CAT}	0.382	0.005*
HH ₁ MOS _{DTAC} ≠ MOS _{CAT}		
HI ₀ MOS _{TRUE} = MOS _{TOT}	0.244	0.786
HI ₁ MOS _{TRUE} ≠ MOS _{TOT}		
HJ ₀ MOS _{TRUE} = MOS _{CAT}	0.258	0.413
HJ ₁ MOS _{TRUE} ≠ MOS _{CAT}		
HK ₀ MOS _{TOT} = MOS _{CAT}	0.026*	0.265
HK ₁ MOS _{TOT} ≠ MOS _{CAT}		

ตารางที่ 5: แสดงค่า P-Value ของสไกป์และเฟสบุ๊ค

T-Test	P-Value
AIS-SK vs AIS-FB	0.181
DTAC-SK vs DTAC-FB	0.324
TRUE-SK vs TRUE-FB	0.451
TOT-SK vs TOT-FB	0.897
CAT-SK vs CAT-FB	0.327

เอกสารอ้างอิง

- [1] “การสื่อสารด้วยระบบ VoIP,” [บทความออนไลน์] แหล่งที่มา <http://www.et.prm.chula.ac.th/pdf/voip1.pdf> [15 กุมภาพันธ์ 2557].
- [2] วรพล ลีลาเกียรติสกุล, “ปัญหาของวีโอไอพี และสแปมบนระบบโทรศัพท์บนอินเทอร์เน็ต,” *Journal of Information Science and Technology* Vol. 2, Issue 1, หน้า 21-30, Jan-Jun 2011.
- [3] Annex B Speech files, ITUT Recommendation P.501. Available from: <http://www.itu.int/net/itu-t/sigdb/genaudi o /AudioForm.g.aspx?val=10 000501>.
- [4] The Phuket News, “Bangkok has ‘most Facebook users’ inworld,” Available from: <http://www.thephuketnews.com/bangkok-has-most-facebook-users-inworld30721.php> [Accessed Jan. 2014].
- [5] Avaya Labs, “Avaya IP Voice Quality Network Requirement,” White paper, Issue3. 2, Aug 2009. Available from: <http://downloads.avaya.com/css/P8/documents/100018203> [Accessed Jan.2014].
- [6] K. Kumaravel, “Comparative Study of 3G and 4G in Mobile Technology,” *IJCSI*, Vol. 8, Issue 5, No. 3, pp. 256-263, Sep 2011.
- [7] L. Lou and X. Xu. “Performance Analysis of a Novel LBS Application Using MBMS & TPEG in 3G Mobile Networks,” *Journal of Networks*, Vol. 7, No. 4, Apr. 2012, pp. 699-706.
- [8] M.Z. Shafiq, J. Lusheng, A.X. Liu, and J. Pang. “Characterizing Geospatial Dynamics of Application Usage in a 3G Cellular Data Network,” *Proc. IEEE INFOCOM*, Orlando, FL, pp. 1341-1349, Mar. 2012.
- [9] Techcrunch, “Facebook Announces Monthly Active Users Were At 1.01 Billion As Of September 30th, An Increase Of 26% Year-Over-Year,” Available from : <http://techcrunch.com/2012/10/23/facebook-announces-monthly-active-users-were-at1-01-billion-as-ofseptember-30th/> [Accessed Jan. 2014].
- [10] M.Z. Shafiq, J. Lusheng, A.X. Liu, and J. Pang, “Characterizing Geospatial Dynamics of Application Usage in a 3G Cellular Data Network,” *Proc. IEEE INFOCOM*, Orlando, FL, pp. 1341-1349, Mar. 2012.
- [11] NEC, “NEC to Launch World’s First, Advanced New Platform for Mobile Operators Employing PICMG Forum Advanced -TCA(TM).” Available from: <http://www.nec.co.jp/press/en/0309/1001.html> [Accessed Jan. 2014].
- [12] “Skype,” [บทความออนไลน์] แหล่งที่มา: <http://www.hitechsky.com/skype/> [15 กุมภาพันธ์ 2557]
- [13] “Facebook” [บทความออนไลน์] แหล่งที่มา: www.thaiail.com/facebook/ [15 กุมภาพันธ์ 2557].
- [14] “การวัดคุณภาพเสียงจากเครือข่ายสื่อสาร,” [บทความออนไลน์] แหล่งที่มา: http://www.tmi.or.th/index.php?option=com_content&view=article&id=293&catid=27:it-infrastructure&Itemid=49 [15 กุมภาพันธ์ 2557].
- [15] ศุภชัย คลังทอง และคณะ, “กรณีศึกษาการเชื่อมต่อระหว่างโครงข่ายโทรศัพท์ผ่านไอพีด้วยเลขหมายในหมวด 06 กับโครงข่ายโทรศัพท์พื้นฐานและโครงข่ายโทรศัพท์เคลื่อนที่ในประเทศไทย,” *IT Innovation for Global Awareness(NCIT 2010)*, กรุงเทพฯ, หน้า 235-239, 28-29 ตุลาคม 2553.
- [16] L. Sun and E. Ifeachor, “Case study of PESQ performance in live wireless mobileVoIP environment,” *Personal, Indoor and Mobile Radio Communications, 2008. PIMRC 2008. IEEE 19th International Symposium on*, pp. 1-6.
- [17] K. Yochanang, T. Daengsi, T. Triyason and P. Wuttidittachotti, “A Comparative Study of VoIP Quality Measurement from G.711 and G.729 Using PESQ and Thai Speech,” *Advances in Information Technology*. Springer International Publishing, pp.242-255, 2013.
- [18] B. Sat and B. Wah, “Evaluation of Conversational Voice Communication Quality of The Skype, Google-Talk, Windows Live and Yahoo Messenger VOIP Systems,” *Multimedia Signal Processing,2007 .MMSp2007.IEEE 9th Workshop on*, pp.135-138.
- [19] Z. Jianming and Y. Dacheng, “Quality of VoIP in Wireless Communication Networks,” *Wireless Communications and Networking Conference WCNC IEEE*, pp. 189-191, 2006.
- [20] B. Rodrigo, et al. “Performance Evaluation of P2P VoIP Application,” *ACM NOSSDAV*. Vol. 7. 2007.
- [21] กสทช., “ผลการวัดคุณภาพการให้บริการโทรศัพท์เคลื่อนที่ วันที่ 22 กันยายน พ.ศ. 2556” [บทความออนไลน์] แหล่งที่มา https://www.dropbox.com/s/scuuenbhcrd6os/2013-09-22_NBTC_VOICE_FTPDL.pdf
- [22] นฤมล ชุมชช และคณะ, “การศึกษาเปรียบเทียบวีโอไอพีด้วยวิธีการวัดแบบพีอีเอสคิว: กรณีศึกษาแอปพลิเคชันไลน์กัยแทงโก้ ผ่านเครือข่ายของ มจพ. และเครือข่าย 3G,” *การประชุมวิชาการทางคอมพิวเตอร์และเทคโนโลยีสารสนเทศ และการประชุมเชิงปฏิบัติการด้านการบริหารและจัดการเครือข่ายยูนิเน็ต สระแก้ว*, หน้า 70-75, มกราคม 2557.
- [23] VassarStats. Concepts & Applications of Inferential Statistics. [online] 2012. [cited 2014 Feb 28]. Available from : <http://vassarstats.net/textbook/index.html>.
- [24] T. Overath, and M. Whalley, t- and F-tests: Testing hypotheses. [online] 2014. [cited 2014 Feb 28]. Available from : [http://www.fil.ion.ucl.ac.uk/spm/doc/mfd/2005/Ft-tests.ppt#280,4,Types of Error](http://www.fil.ion.ucl.ac.uk/spm/doc/mfd/2005/Ft-tests.ppt#280,4,Types%20of%20Error).

การวิเคราะห์การเชื่อมต่อระบบเครือข่ายคอมพิวเตอร์ โดยทฤษฎีกราฟ

กรณีศึกษาวิทยาลัยพาณิชย์การธนบุรี

Analysis of Connecting Computer Network System with Graph Theory A Case Study of Thonburi Commercial College

สุภาพร เกิดกิจ (Supaporn Kurdkit)¹ ล้วนกร สร้อยมาต (Lawankorn Soimart)² และ

สุนันทา สดสี (Sunantha Sodsee)³

¹ คณะคอมพิวเตอร์ธุรกิจ วิทยาลัยพาณิชย์การธนบุรี

² คณะวิทยาการจัดการ มหาวิทยาลัยราชภัฏนครสวรรค์

³ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹chama2005@gmail.com, ²lawankorn.s@gmail.com, ³sunanthas@kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาพฤติกรรมการเชื่อมต่อของเครือข่ายคอมพิวเตอร์ กรณีศึกษาวิทยาลัยพาณิชย์การธนบุรี เพื่อดูความสัมพันธ์ระหว่างเครื่องคอมพิวเตอร์ โดยใช้การวิเคราะห์คุณสมบัติเครือข่ายด้วยทฤษฎีกราฟ (Graph Theory) เช่น Degree Centrality (DC) และ Betweenness Centrality (BC) ในการสร้างแบบจำลองเครือข่ายและวิเคราะห์เครือข่าย ผลการวิเคราะห์แสดงให้เห็นว่าห้องควบคุมระบบเครือข่ายคอมพิวเตอร์ อาคาร 7 และห้องปฏิบัติการอาคาร 8 เป็นจุดเชื่อมต่อที่สำคัญโดยพิจารณาจากห้องควบคุมระบบเครือข่ายคอมพิวเตอร์อาคาร 7 มีค่า $DC = 89$ และค่า $BC = 0.7974$ ห้องปฏิบัติการคอมพิวเตอร์อาคาร 8 มีค่า $DC = 202$ และค่า $BC = 0.6107$ ดังนั้น อาคาร 7 และอาคาร 8 จึงเป็นจุดเชื่อมต่อที่สำคัญที่สุด เนื่องจากเป็นจุดศูนย์กลางและจุดเชื่อม

คำสำคัญ : เครือข่ายซับซ้อน ทฤษฎีกราฟ เครือข่ายคอมพิวเตอร์ ค่ากลาง (Centrality)

Abstract

The purpose of this paper is to study the behavior of a computer connected to a network in relation to buildings. This investigation has yielded significant information referring to computer. Thonburi Commercial College was chosen for the experimental testing, analyzing properties of network with graph theory. Such as the Degree and Betweenness Centrality were also applied to

evaluate the prototype network and analysis network. The results showed that the server control room at building 7 and lab room at building 8, which is at the significant connection point. It was found that 1) Value of Degree Centrality measurement of maximum links connection of the server control room at building 7 = 89, value of Betweenness Centrality measurement of maximum links connections at building 7 = 0.7947. 2) Value of Degree Centrality measurement of maximum links connection of the lab room at building 8 = 202, value of Betweenness Centrality measurement of maximum links connections at building 8 = 0.6107. In conclusion, building 7 and building 8 were found to be of importance due to being located at the center point

Keywords : Complex Network, Graph Theory, Computer Network, Centrality

1. บทนำ

วิทยาลัยพาณิชย์การธนบุรี ได้ติดตั้งระบบเครือข่ายคอมพิวเตอร์โดยเชื่อมโยงเครือข่ายไปยังอาคารต่าง ๆ ภายในบริเวณวิทยาลัยฯ ปัจจุบันมีการใช้งานเครือข่ายเพิ่มขึ้นกว่า 500 เครื่อง [1] และมีแนวโน้มที่จะเพิ่มจำนวนเครื่องมากขึ้น ส่งผลต่อการจัดการดูแลระบบเครือข่าย เนื่องจากปัจจุบันมีการเก็บข้อมูลการใช้งานเครือข่ายในรูปแบบฐานข้อมูลและใช้โปรแกรมประยุกต์ในการจัดการเครือข่าย [1] เมื่อเครือข่ายมีการเชื่อมโยงกันซับซ้อนมากขึ้น จึงต้องนำโปรแกรมที่มี

คุณสมบัติในการวิเคราะห์เชิงสถิติ มาช่วยอำนวยความสะดวกให้สามารถจัดการบริหารเครือข่ายได้อย่างเหมาะสมและมีประสิทธิภาพ ในกรณีระบบเครือข่ายเกิดความเสียหาย หรือปัญหาไวรัส เช่น การวิเคราะห์หาจุดเชื่อมต่อเครือข่ายที่มีความสำคัญ การยุติการเชื่อมต่อชั่วคราวเมื่อเกิดการติดไวรัสในระบบที่จุดใดในเครือข่าย ผู้วิจัยจึงมีแนวคิดในการวิเคราะห์พฤติกรรมของการเชื่อมต่อระบบเครือข่ายของวิทยาลัยฯ เพื่อหาอุปกรณ์กระจายสัญญาณที่มีความสำคัญภายในระบบเครือข่ายสำหรับการวิเคราะห์ระบบเครือข่ายคอมพิวเตอร์ครั้งนี้ ผู้วิจัยมีการประยุกต์ใช้ทฤษฎีกราฟ [2] ในการจำลองเป็นแผนภาพแสดงความสัมพันธ์ระหว่างอุปกรณ์เชื่อมต่อ โดยใช้โปรแกรม Pajek [3] ในการวิเคราะห์คุณลักษณะค่า Degree Centrality (DC) ค่า Betweenness Centrality (BC) ค่า DC พิจารณาจากจำนวนเส้นเชื่อมต่ออุปกรณ์ที่อยู่ภายในอาคารเดียวกันและอาคารอื่น ๆ ที่มีปริมาณสูงสุด คำนวณจากเส้นเชื่อมต่อภายในหรือภายนอกอาคาร ค่า BC พิจารณาจากค่าคั่นกลางที่มีค่าสูงสุดและระยะทางสั้นสุด ผลที่ได้จากการศึกษาจะทำให้สามารถนำข้อมูลดังกล่าวไปช่วยในการตัดสินใจเพื่อปรับปรุงวางแผน ออกแบบระบบเครือข่ายคอมพิวเตอร์ที่จะขยายตัวเพิ่มขึ้นในอนาคต ก่อให้เกิดประสิทธิภาพในการให้บริการ และใช้งานทรัพยากรร่วมกันอย่างสูงสุด ในงานวิจัยนี้แบ่งเนื้อหา ดังนี้ ส่วนที่ 2 ทฤษฎีที่เกี่ยวข้อง ส่วนที่ 3 วิธีดำเนินงาน ส่วนที่ 4 ผลการดำเนินงาน และส่วนที่ 5 บทสรุปและข้อเสนอแนะ

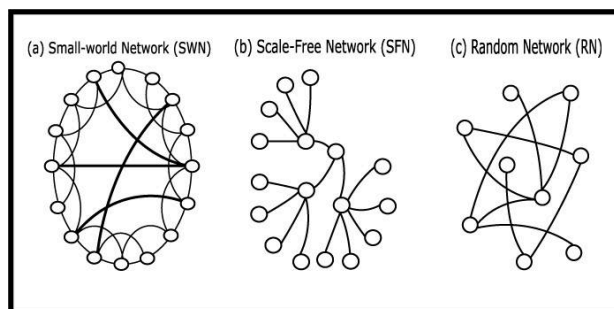
2. ทฤษฎีที่เกี่ยวข้อง

การวิเคราะห์การเชื่อมโยงระบบเครือข่ายคอมพิวเตอร์ โดยทฤษฎีกราฟ กรณีศึกษาวิทยาลัยพัฒนวิชาการธนบุรี มีทฤษฎีที่เกี่ยวข้อง ดังนี้

2.1 ทฤษฎีเครือข่ายซับซ้อน (Complex Network)

Complex Network [4] มีลักษณะการเชื่อมต่อเครือข่ายที่ไม่มีการเพิ่มจำนวนหรือลดจำนวนของโหนดและเส้นเชื่อมต่อกันอยู่ตลอดเวลา มักเป็นเครือข่ายที่มีขนาดใหญ่ ซับซ้อน ลักษณะการเชื่อมต่อเลียนแบบพฤติกรรมเครือข่ายในธรรมชาติ เช่น ความสัมพันธ์ระหว่างมนุษย์ทางสังคม ธุรกิจ รวมทั้งพฤติกรรมของสัตว์บางชนิดที่อาศัยร่วมกันเป็นฝูง ส่วนประกอบทาง

โครงสร้างของ Complex Network ได้แก่ โหนด (Node) และเส้นเชื่อมโยง (Edge) เชื่อมต่อเข้าด้วยกันก่อให้เกิดเครือข่ายที่มีขนาดใหญ่และซับซ้อน ในการค้นพบรูปแบบการเชื่อมต่อแบ่งเป็น 3 ประเภท คือ Random Network Model, [5] Small-World Network Model [6] และ Scale-Free Network Model [7] แสดงดังภาพที่ 1



ภาพที่ 1: รูปแบบ Complex Network ทั้ง 3 รูปแบบ [8]

2.1.1 Random Network Model

เป็นการเชื่อมต่อกันระหว่างโหนดจากการเพิ่มโหนดที่เป็นสมาชิกของเครือข่ายเพื่อเชื่อมต่อดังวิธีการสุ่ม ทั้งนี้ขึ้นอยู่กับจำนวนโหนดและจำนวนเส้นเชื่อมโยง

2.1.2 Small-World Network Model

เกิดจากกลุ่มเครือข่าย (Cluster) หลายกลุ่มเครือข่ายเชื่อมต่อเข้าด้วยกัน ลักษณะความสัมพันธ์ของการเชื่อมโยง มองถึงโหนดที่เชื่อมอยู่ใกล้กัน (Local) จะมีความสัมพันธ์กัน มากกว่าโหนดที่อยู่ไกลกัน

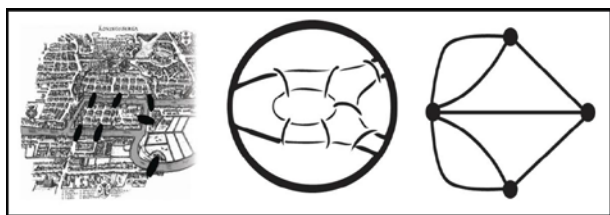
2.1.3 Scale- Free Network Model

รูปแบบการเชื่อมต่อใช้กฎการกระจายตัวแบบ Power Law Degree Distribution กล่าวคือจำนวนโหนดที่มีการเชื่อมต่อในปริมาณที่มากจะมีจำนวนโหนดที่มีอยู่น้อย นิยมเรียกว่า “Popular Node” ส่วนจำนวนโหนดที่มีการเชื่อมต่อในปริมาณที่น้อยจะมีจำนวนโหนดที่มีอยู่มาก

2.2 ทฤษฎีกราฟ (Graph Theory)

ทฤษฎีกราฟมีต้นกำเนิดมาจากการแก้ปัญหาสะพานทั้ง 7 รู้จักกันในนาม “ปัญหาสะพานทั้งเจ็ดแห่งเมืองเคอนิกส์แบร์ก” การแก้ปัญหาจะทำได้ก็ต่อเมื่อกราฟนั้นไม่มีจุดยอดที่มีเส้นเชื่อมเป็นจำนวนคี่ ซึ่งได้มีการนำรูปแบบการเชื่อมต่อ (Topology) มาประยุกต์ใช้กับทฤษฎีกราฟและการเชื่อมต่อ

ที่สนใจเฉพาะ โครงสร้างรูปร่างชนิดที่ไม่ขึ้นกับขนาด ระยะ หรือการวัด [9] แสดงในภาพที่ 2



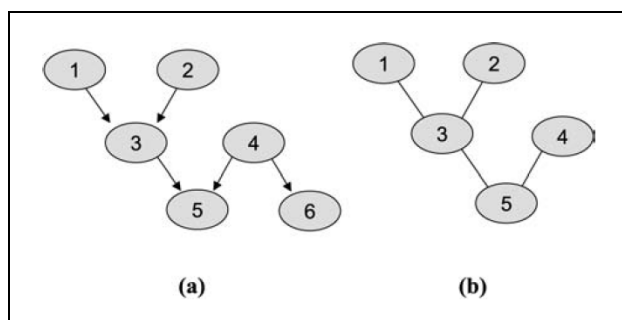
ภาพที่ 2: สะพานทั้งเจ็ดแห่งเมืองเคอนิกส์แบร์ก [9]

ทฤษฎีกราฟมักถูกนำเสนอในลักษณะของรูปภาพ โดยใช้จุดแทนจุดยอดของแต่ละจุด และลากเส้นระหว่างจุดยอดถ้าจุดยอดทั้งสองนั้นมีเส้นเชื่อมถึงกัน ในกรณีกราฟมีทิศทาง ทิศทางของเส้นเชื่อมจะถูกระบุโดยใช้ลูกศรในการอ้างอิงทิศทางนั้น ๆ สามารถเขียนนิยามของกราฟได้ตามสมการดังนี้ [9]

$$G = (V_G, E_G) \tag{1}$$

โดยที่ $E_G \subseteq E_G(V_G)$ ให้ V_G แทนเซตของจุดในกราฟ G และให้ E_G แทนเซตของเส้นเชื่อมต่อระหว่างจุดในกราฟ G

การจำแนกชนิดของกราฟ แบ่งได้ตามลักษณะข้อมูลที่เก็บได้แก่ กราฟแบบมีทิศทาง (Directed Graph) คือกราฟที่มีลูกศรกำกับเส้นเพื่อบอกจำนวนครั้งที่เข้ามาและจำนวนครั้งที่ออกไป ส่วนกราฟแบบไม่มีทิศทาง (Undirected Graph) คือกราฟที่ไม่มีลูกศรกำกับเส้น [10] แสดงในภาพที่ 3



ภาพที่ 3: (a) กราฟแบบมีทิศทางและ (b) กราฟแบบไม่มีทิศทาง

2.2 การวัดคุณสมบัติโดยทฤษฎีกราฟ

หลักการเบื้องต้นของวิธีการวิเคราะห์เครือข่าย วัดคุณสมบัติ ดังนี้

2.2.1 ค่าความเป็นจุดศูนย์กลาง (Centrality) [11]

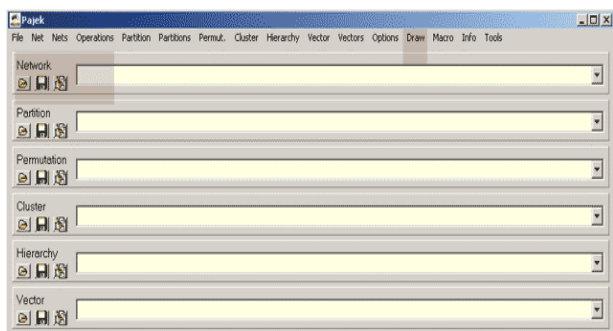
เป็นการวัดค่าความเป็นจุดศูนย์กลางของสมาชิกเครือข่าย แต่ละจุด หรือวัดค่าความเป็นศูนย์กลางของเครือข่าย ซึ่งวัดจากจำนวนสมาชิกทั้งหมดในเครือข่าย วิธีการวัดค่าความเป็นจุดศูนย์กลางของสมาชิกเครือข่าย ได้แก่

1) ความเป็นจุดศูนย์กลางโดยวัดจากระดับ (Degree Centrality: DC) [11] เป็นการค้นหาว่าสมาชิกเครือข่ายใดที่เป็นจุดศูนย์กลางของการเชื่อมโยง ซึ่งเป็นตำแหน่งที่มีอิทธิพลสูงสุดในเครือข่าย วัดได้จากจำนวนเส้นเชื่อมโยงทั้งหมดที่โยงมาจากสมาชิกเครือข่ายอื่น ๆ สมาชิกนั้นอาจอยู่ภายในกลุ่มเดียวกันและข้ามกลุ่ม

2) ความเป็นจุดศูนย์กลางโดยวัดจากการค้นกลาง (Betweenness Centrality: BC) [11] เป็นการค้นหาว่าสมาชิกเครือข่ายใดที่มีตำแหน่งเป็นสะพาน (Bridges) เชื่อมกลุ่มเครือข่ายต่าง ๆ ที่อยู่ห่างกันให้เข้าหากัน ทำหน้าที่เป็นตัวกลางในการติดต่อเชื่อมโยงระหว่างสมาชิกเครือข่ายอื่น ๆ มีบทบาทสำคัญในการเปลี่ยนแปลงรูปแบบการเชื่อมต่อเครือข่าย รวมทั้งผลกระทบต่าง ๆ เข้ามาในเครือข่าย ใช้วิธีคำนวณจากสัดส่วนของระยะทางที่สั้นที่สุด (Geodesic path) ในการเชื่อมต่อระหว่างสมาชิกเครือข่ายแต่ละคู่ [11]

2.3 ทฤษฎี Pajek

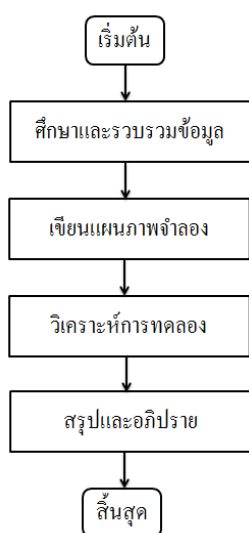
Pajek เป็นโปรแกรมที่ไม่เสียค่าใช้จ่าย สามารถรองรับการประมวลเครือข่ายที่มีขนาดใหญ่และซับซ้อน วิเคราะห์ผลเชิงสถิติ (Statistic) และแสดงแผนผังเครือข่าย (Visualization) นอกจากนี้สามารถนำมาใช้ในการวิเคราะห์เครือข่าย ซึ่งจะประกอบไปด้วยกราฟและข้อมูลอยู่ในรูปแบบของโหนดและเส้นเชื่อมต่อ การประมวลผลจะอยู่ในรูปแบบของ Network Data File สามารถสร้างได้จากกลุ่มโปรแกรม Editor หรือประยุกต์ใช้ Macro เพื่อ Run บน Command Line ได้ นอกจากนี้โปรแกรม Pajek มีความยืดหยุ่นในการเพิ่มข้อมูล แก้ไข การเลือกประเภทการวิเคราะห์ให้เหมาะสมกับลักษณะของงาน ปรับแต่งการแสดงผล แสดงผลบนจอภาพและส่งออกข้อมูล และภาพ โดยบันทึกเป็นแฟ้มข้อมูลและนำไปใช้รายงานได้ [11] การใช้โปรแกรม Pajek แสดงดังภาพที่ 4



ภาพที่ 4: โปรแกรม Pajek [12]

3. วิธีการดำเนินการวิจัย

วิธีดำเนินงานวิจัยสำหรับการจำลองและวิเคราะห์ การเชื่อมโยงระบบเครือข่ายคอมพิวเตอร์ มีกระบวนการดำเนินงาน แสดงดังภาพที่ 5



ภาพที่ 5: วิธีดำเนินการ

3.1 ศึกษาและรวบรวมข้อมูล

ผู้ศึกษาได้กำหนดข้อมูลในการศึกษาออกเป็น 2 ส่วน ส่วนที่หนึ่งคือ ข้อมูลจำนวนเครื่องคอมพิวเตอร์และจำนวน อุปกรณ์กระจายสัญญาณภายในเครือข่ายวิทยาลัยฯ ส่วนที่สองคือข้อมูลการเชื่อมโยงระหว่างอาคารภายในเครือข่ายวิทยาลัยฯ จากการศึกษารวบรวมข้อมูล พบว่าในส่วนที่ หนึ่ง จำนวนเครื่องคอมพิวเตอร์ที่เชื่อมต่อตามอาคารต่าง ๆ มีจำนวนทั้งหมด 536 เครื่อง และอุปกรณ์กระจายสัญญาณมีจำนวนทั้งหมด 7 ตัว แบ่งเป็นอุปกรณ์กระจายสัญญาณประเภท

L3 จำนวน 1 ตัว อุปกรณ์กระจายสัญญาณประเภท L2 จำนวน 6 ตัว และ Proxy Server 1 เครื่อง ในส่วนที่สอง การเชื่อมโยงระหว่างอาคารต่าง ๆ มีจำนวน 7 อาคาร ประกอบด้วย Building 2, Building 3, Building 5, Building 6, Building 7, Building 8, Building 9

3.2 จำลองเครือข่ายคอมพิวเตอร์ด้วยกราฟ

ในการสร้างภาพกราฟ ผู้วิจัยได้กำหนดว่าทุกจุดที่เชื่อมต่อกันในระบบเครือข่ายมีสถานะการเชื่อมต่อเท่าเทียมกันทั้งหมด และเป็นการเชื่อมต่อแบบไม่มีทิศทาง จากนั้นใช้โปรแกรม Pajek สร้างกราฟ เลือกประเภท Partition เพื่อทำการแยกกลุ่มของอุปกรณ์ ได้แก่ Switch, Proxy Server และ Client เนื่องจากเป็นการวิเคราะห์การเชื่อมต่อภายในบริเวณวิทยาลัยฯ ส่วนประเภท Vector นำมาใช้ในการวิเคราะห์ค่า DC และ BC

3.3 วิเคราะห์คุณลักษณะเครือข่ายคอมพิวเตอร์

การศึกษาครั้งนี้จะวิเคราะห์คุณลักษณะการเชื่อมต่ออุปกรณ์คอมพิวเตอร์ เพื่อคำนวณหาค่าความเป็นศูนย์กลางของระบบเครือข่ายคอมพิวเตอร์ ดังนี้

1) ค่า Degree Centrality (DC) โดยวัดจากจำนวนหรือปริมาณ โดยมีสมการดังนี้

$$d(i) = \sum_j m_{ij} \tag{2}$$

โดยที่ $m_{ij} = 1$ ถ้ามีการเชื่อมต่อระหว่างโหนด i และ j และ หรือ $m_{ij} = 0$ ถ้าไม่มีการเชื่อมต่อระหว่างโหนด i และ j

2) ค่า Betweenness Centrality (BC) โดยมีสมการดังนี้

$$BC_u = \sum \frac{\sum_{i \in S_{ij}} \delta_i^u}{|S_{ij}|} \tag{3}$$

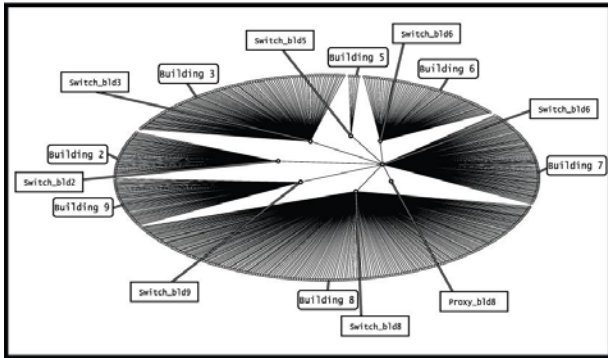
ในที่นี้ BC_u แสดงถึงค่า BC ของโหนด u โดยให้ $i \in S_{ij} \delta_i^u$ แทนจำนวนระยะทางที่สั้นที่สุดจากโหนด i ไปโหนด j ที่ผ่านจุด δ_i^u และให้ S_{ij} แทนจำนวนระยะทางที่สั้นที่สุดจากโหนด i ไป โหนด j ซึ่ง $i \neq j$

4. ผลการดำเนินงาน

ในส่วนนี้จะแสดงผลการวิเคราะห์คุณลักษณะเครือข่ายคอมพิวเตอร์

4.1 ผลการจำลองระบบเครือข่ายคอมพิวเตอร์

จำลองแผนภาพได้ตามภาพที่ 6

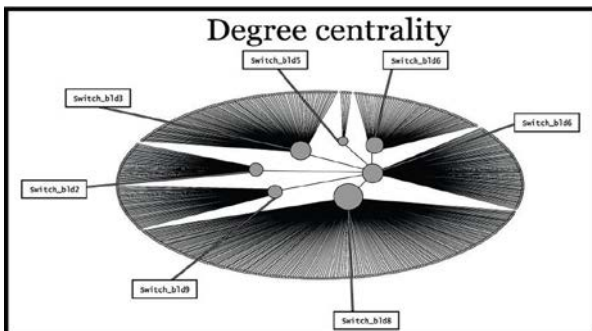


ภาพที่ 6: แผนภาพจำลองระบบเครือข่ายคอมพิวเตอร์

ผลการสร้างแบบจำลองระบบเครือข่ายคอมพิวเตอร์ ได้แผนภาพการเชื่อมต่อแสดงดังภาพที่ 6 ประกอบด้วย Switch, Proxy Server และเครื่องคอมพิวเตอร์ที่อยู่ตามห้องปฏิบัติการตามอาคารต่าง ๆ ซึ่ง Switch และ Proxy Server ที่อาคาร 7 มีปริมาณการเชื่อมต่อมากที่สุด

4.2 ผลการวิเคราะห์เครือข่ายคอมพิวเตอร์

ผลการวิเคราะห์ระบบเครือข่ายคอมพิวเตอร์ แสดงให้เห็นว่าขนาดของจุดที่อาคาร 8 ใหญ่ที่สุด มีค่า DC = 202 สูงสุดเป็นลำดับแรก และอาคาร 7 มีค่า DC = 89 รองเป็นอันดับ 3 ส่วนลำดับที่ 5 ที่อาคาร 9 มีค่า DC = 42 นำเสนอดังภาพที่ 5 ส่วนอาคาร 7 มีค่า BC = 0.7974 สูงสุดเป็นอันดับแรก และอาคาร 8 มีค่า BC = 0.6107 รองเป็นลำดับที่ 2 ส่วนลำดับที่ 5 ที่อาคาร 9 มีค่า BC = 0.1475 นำเสนอดังภาพที่ 7

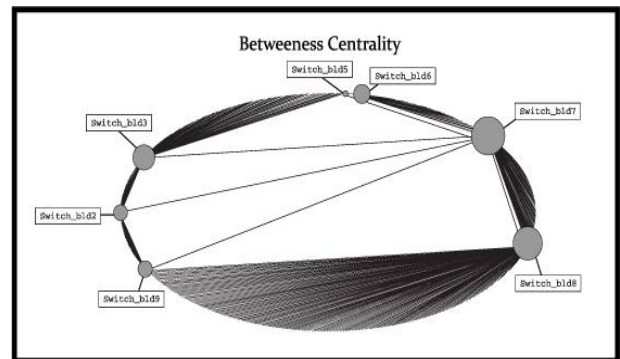


ภาพที่ 7: แบบจำลอง Degree Centrality

ตารางที่ 1: ค่า Degree Centrality ที่ได้จากการทดลอง

5 อันดับแรกสูงสุด

Rank	Vertex	Value	Id
1	293	202.00	Switch_bld8
2	42	101.00	Switch_bld3
3	211	89.00	Switch_bld7
4	149	62.00	Switch_bld6
5	495	42.00	Switch_bld9



ภาพที่ 8: แบบจำลอง Betweenness Centrality

ตารางที่ 2: ค่า Betweenness Centrality ที่ได้จากการทดลอง

5 อันดับแรกสูงสุด

Rank	Vertex	Value	Id
1	211	0.7974	Switch_bld7
2	293	0.6107	Switch_bld8
3	42	0.3392	Switch_bld3
4	149	0.2152	Switch_bld6
5	495	0.1475	Switch_bld9

จากผลการวิเคราะห์ข้างต้น อาคาร 8 มีค่า DC เป็นลำดับแรก และ BC มีค่าเป็นลำดับที่ 2 ส่วนอาคาร 7 มีค่า BC สูงที่สุดเป็นลำดับแรก แต่ค่า DC มีค่าสูงอยู่ที่ลำดับที่ 3 แสดงตามตารางที่ 1 และตารางที่ 2

5. บทสรุป

ในที่นี้เมื่อพิจารณาค่า DC และ BC แล้วสรุปได้ว่าห้องควบคุมเครือข่ายคอมพิวเตอร์อาคาร 7 มีค่า DC = 89 และค่า BC = 0.7974 ห้องปฏิบัติการคอมพิวเตอร์อาคาร 8 มีค่า

DC = 202 และค่า BC = 0.6107 ซึ่งอาคาร 8 เป็นศูนย์กลางในการเชื่อมต่อเนื่องจากค่า DC และ BC มีค่าที่สูงทั้งคู่

เนื่องจากการวัดค่ากลาง ในงานวิจัยนี้ใช้การวัดค่า DC และค่า BC เท่านั้น ควรใช้ค่ากลางอื่นในการวัดเพื่อวิเคราะห์คุณลักษณะของเครือข่ายคอมพิวเตอร์ เช่น Closeness Centrality (CC) และ Eigenvector Centrality (EC) เพื่อเพิ่มประสิทธิภาพการวิเคราะห์คุณสมบัติเครือข่ายของวิทยาลัยฯ นอกจากนี้ควรเพิ่มการวิเคราะห์คุณลักษณะโครงสร้างลำดับชั้น (Hierarchical Network) เพื่อค้นหาตัวแปรที่สำคัญ เพื่อให้ได้ความใกล้เคียงกับของจริงมากที่สุด

เอกสารอ้างอิง

- [1] สุภาพร เกิดกิจ, "TCC Diagram Networks," <http://panitthon.ac.th/diagram>, 12 Feb 2014.
- [2] A. Gibbons, *Algorithmic Graph Theory*, The Press Syndicate of the University of Cambridge, 1985.
- [3] R. Wittayawuttikul and N. Wipawin, "Social Network Analysis (SNA) for Information Science Research," *Journal of Library and Information Science Srinakharinwirot University*, vol. 2012, no. 2, Dec. 2012, pp. 127–139.
- [4] A.-L. Barabási and E. Bonabeau, "Scale-free networks," *Science*, vol. 288, pp. 60–69, 2003.
- [5] A.-L. Barabási and R. Albert, "Emergence of scaling in Random Networks," *Science*, vol. 286, 1999, pp. 509–512.
- [6] D. J. Watts and S. H. Strogatz, "Collective dynamics of small-world networks," *nature*, vol. 393, pp. 440–442, 1998.
- [7] A.-L. Barabási and E. Bonabeau, "Scale-free networks," *Science*, vol. 288, pp. 60–69, 2003.
- [8] R. Diestel, *Graph Theory*, Springer-erlag:Heidelberg, 2010.
- [9] K. Rosen, "Teaching Mathematics Applications," vol. 2010, no. 3, Apr. 2010, pp. 155–163.
- [10] M.F. Renzi, P. Vicard, and R. Guglielmetti, "Probabilistic expert systems for managing information to improve services," *The TQM Journal*, vol. 2009, no. 4, pp. 429–442.
- [11] S.Tiyawongsuwan, "Social Network and Urban Regeneration Processes: A Case Study of Ratchanikunl 3 Community, Thailand.," *Proc. 3rd Int. Conf. Asian Assoc. Urban Reg. Stud. AAURS Top. Valuab. Risky Better Qual. Life Environ. Sustain. Plan. Gold. Tulip Sover. Hotel Bangk.*, vol. 2011, Dec. 2011, pp. 124–135.
- [12] "Pajek Tutorial," <http://iv.slis.indiana.edu/lm/lm-pajek.html>, 14 Feb 2014.

การวิเคราะห์เครือข่ายทางสังคมโดยทฤษฎีกราฟ :
กรณีศึกษา คณาจารย์คณะบริหารธุรกิจ มหาวิทยาลัยกรุงเทพธนบุรี
Analysis of Social Networking by Graph Theory: The Case of Faculty Members
of Business Administration, Bangkok Thonburi University

ชนนิกานต์ รอดมรณ์ (Chonnikarn Rodmorn)¹ และ มธุรส ผ่านเมือง (Mathuros Panmuang)²

^{1,2} สาขาวิชาคอมพิวเตอร์ธุรกิจ คณะบริหารธุรกิจ มหาวิทยาลัยกรุงเทพธนบุรี

¹ r.chonnikarn@hotmail.com, ² mathuros.p@hotmail.com

บทคัดย่อ

การวิเคราะห์เครือข่ายทางสังคมคือการแสดงความสัมพันธ์ของโครงสร้างทางสังคม โดยเครือข่ายทางสังคมสามารถอธิบายลักษณะความสัมพันธ์ทางสังคมระหว่างบุคคลหรือสิ่งต่าง ๆ ที่อยู่รอบตัวเรา ทำให้เกิดความรู้ความเข้าใจถึงลักษณะโครงสร้างของความสัมพันธ์และปฏิสัมพันธ์ระหว่างสมาชิกและการไหลเวียนของสารสนเทศ ในการศึกษาครั้งนี้ได้กำหนดให้โหนดแสดงถึงความเป็นตัวแทนของคณาจารย์คณะบริหารธุรกิจ และลิงค์คือความสัมพันธ์ของโหนดที่มีความเกี่ยวเนื่องกัน โดยการศึกษาครั้งนี้ได้วิเคราะห์หาค่าความเป็นศูนย์กลางของโครงข่ายทางสังคมของคณาจารย์คณะบริหารธุรกิจ ได้แก่ 1) Degree Centrality ได้ค่าสูงสุดคือ 76 2) Closeness Centrality ได้ค่าสูงสุดคือ 0.602 และ 3) Betweenness Centrality ได้ค่าสูงสุดคือ 0.163 ซึ่งจากค่าสูงสุดทั้งสามค่า พบว่า อาจารย์ IBM คือบุคคลที่สำคัญที่สุดในเครือข่าย กล่าวคือ เมื่อต้องการส่งข้อมูลข่าวสารไปยังคณาจารย์ภายในคณะบริหารธุรกิจ ควรพิจารณาอาจารย์ IBM ให้เป็นผู้กระจายข่าวสารเป็นอันดับแรก เพราะอาจารย์ IBM คือบุคคลที่เป็นตัวกลางในการกระจายข้อมูลไปยังอาจารย์ท่านอื่น ๆ ทำให้คณาจารย์ในคณะบริหารธุรกิจจะได้รับข้อมูลข่าวสารอย่างรวดเร็ว

คำสำคัญ: เครือข่ายทางสังคม การวิเคราะห์เครือข่าย

Abstract

Social networking analysis is the presentation concerning relationships of social structure which is used to explain the relationships between members of social system or surrounded things which would build the knowledge about relationships and flow of information. In this study, Node was represented as the faculty members of business administration and Link was the relevant relationship of Node. The study aims to find centrality value of networking including 1) Degree Centrality with the highest value equal to 76, 2) Closeness Centrality with the highest value equal to 0.602, and 3) Betweenness Centrality with the highest value equal to 0.163. From the top three highest values, it was found that lecturer IBM was the most important person in the networking. That is, when it is required to send the information to the faculty members of business administration, it should be considered that the lecturer IBM was the first person to distribute the information. The lecturer IBM was the person who was the center to distribute all information to other faculty members, so all faculty members of business administration were able to receive all information quickly.

Keyword: Social Networking, Network Analysis

1. บทนำ

เครือข่ายทางสังคม (Social network) เป็นเรื่องที่อยู่ในความสนใจในสังคมยุคปัจจุบันและเกี่ยวข้องกับสาขาวิชาต่าง ๆ แทบทุกวงการ ทั้งทางด้านสังคมศาสตร์และวิทยาศาสตร์ [1] เนื่องจากเครือข่ายทางสังคมสามารถอธิบายลักษณะความสัมพันธ์ทางสังคมระหว่างบุคคลหรือสิ่งต่าง ๆ ที่อยู่รอบตัวเรา ทำให้เกิดความรู้ความเข้าใจถึงลักษณะโครงสร้างความสัมพันธ์ระหว่างสมาชิกภายในเครือข่าย เมื่อเกิดเครือข่ายทางสังคมขึ้นเมื่อไหร่ ความสัมพันธ์ระหว่างสมาชิกที่อยู่ในเครือข่ายย่อมเกิดขึ้นตามมา ดังนั้นการทำให้สมาชิกทั้งหมดภายในเครือข่ายสามารถเชื่อมโยงเข้าหากันได้อย่างครบถ้วนย่อมทำให้เกิดการรับส่งข้อมูลที่มีคุณภาพ โดย Newman [2] ได้แบ่งเครือข่ายทางสังคมออกเป็น 4 ประเภท ได้แก่ 1) เครือข่ายทางสังคม เช่น เครือข่ายของบุคคล กลุ่มบุคคล หรือองค์กร เป็นต้น 2) เครือข่ายสารสนเทศ เช่น เครือข่ายเว็ลด์ไวด์เว็บ เครือข่ายการอ้างอิงบทความวิจัย เป็นต้น 3) เครือข่ายทางเทคโนโลยี เช่น เครือข่ายอุปกรณ์คอมพิวเตอร์ เครือข่ายอินเทอร์เน็ต เครือข่ายคอมพิวเตอร์ เครือข่ายโทรศัพท์ เป็นต้น 4) เครือข่ายทางชีววิทยา เช่น เครือข่ายเส้นใยประสาท เครือข่ายสายใยอาหาร เครือข่ายการควบคุมยีน เครือข่ายวิถีเมตาบอลิซึมในเซลล์ของสิ่งมีชีวิต เป็นต้น

เครือข่ายทางสังคมเป็นการเปิดโอกาสให้สมาชิกหรือผู้ใช้เครือข่ายสามารถสร้างข้อมูลส่วนตัวเพื่อเชื่อมโยงไปยังเพื่อนและเพื่อนของเพื่อนได้ นอกจากนั้นเครือข่ายทางสังคมยังเป็นช่องทางในการแบ่งปันข้อมูล การขอความช่วยเหลือ การส่งข้อมูลข่าวสารในภาวะวิกฤติ ซึ่งนำไปสู่การสร้างความสัมพันธ์ของผู้คนมากมายทั่วโลก [3] และในปัจจุบันเกิดการประยุกต์ใช้เครือข่ายทางสังคมในรูปแบบต่าง ๆ เช่น Facebook, MySpace, Twitter, LinkedIn, Instagram เป็นต้น

ดังนั้นในการศึกษาค้นคว้าครั้งนี้ผู้วิจัยจึงได้ศึกษาเครือข่ายทางสังคมในรูปแบบของกลุ่มบุคคล โดยศึกษาความสัมพันธ์ของคณาจารย์คณะบริหารธุรกิจ มหาวิทยาลัยกรุงเทพธนบุรีทั้งนี้ในปัจจุบันคณาจารย์คณะบริหารธุรกิจ มีจำนวนทั้งสิ้น 137 คน โดยแบ่งเป็นอาจารย์ที่ปฏิบัติงานประจำ และอาจารย์ที่ปฏิบัติงานชั่วคราว คือ มีเพียงภาระงานในการสอนหนังสือเท่านั้น โดยไม่มีภาระในการปฏิบัติงานอื่น ๆ ซึ่งจะมา

ปฏิบัติงานเป็นครั้งคราวตามแต่ภาระการสอนที่เกิดขึ้นเท่านั้น นอกจากนี้อาจารย์ในคณะบริหารธุรกิจบางท่านยังมีห้องพักอยู่คนละอาคาร ด้วยเหตุนี้จึงเกิดปัญหาในการติดต่อสื่อสาร กล่าวคือ เมื่อมีภาระงานที่ต้องรับผิดชอบร่วมกันจะไม่สามารถกระจายข่าวสารได้อย่างทั่วถึง จากการทบทวนงานวิจัยที่เกี่ยวข้องพบว่า การวิเคราะห์เครือข่ายทางสังคมนั้นมีผู้ศึกษาอยู่มากมาย เช่น วิทยา เรื่องฤทธิ์ [4] ได้ศึกษาการวิเคราะห์เครือข่ายทางสังคมของกระแสการเดินทางไปทำงานในภูมิภาคเมืองโคราช โดยภาพรวมของเครือข่ายการเดินทางไปทำงานในเมืองภูมิภาคโคราชมีการจัดเรียงตัวในลักษณะเข้าสู่ศูนย์กลางค่อนข้างสูง โดยมีค่า Network Centralization ร้อยละ 78.28 ทั้งนี้อาจเป็นเพราะการสัญจรไปทำงานเป็นเครือข่ายเชิงพื้นที่ที่มีจุดเริ่มต้นและปลายทางชัดเจนจากที่อยู่อาศัยของตนเดินทางออกไปสู่แหล่งการทำงาน ซึ่งมีกระจายตัวอยู่ในตัวเมืองหลักและเมืองต่างๆ ที่มีการพัฒนาลงทุนใหม่ภายในพื้นที่เมืองภูมิภาค นอกจากนั้น ทัตรนนท์ พุ่มนุช [5] ยังได้นำเสนอแนวทางการใช้เครือข่ายสังคมในการปฏิบัติงานของบุคลากรทางการศึกษาคือ 1) ควรมีลักษณะการสื่อสารทั้งกลุ่มเดียวและต่างกลุ่มกัน 2) ควรจัดหาเครื่องมือในเครือข่ายสังคม 3) ควรมีการเผยแพร่และให้ความรู้เกี่ยวกับการใช้เครือข่ายสังคม 4) ควรได้รับการสนับสนุนด้านพัฒนาบุคลากรให้มีความรู้เกี่ยวกับการใช้เครือข่ายทางสังคม

จากการทบทวนวรรณกรรมและงานวิจัยที่เกี่ยวข้อง รวมถึงปัญหาที่เกิดขึ้นในการติดต่อสื่อสารหรือแลกเปลี่ยนข้อมูลระหว่างคณาจารย์ในคณะบริหารธุรกิจ ผู้วิจัยจึงวิเคราะห์เครือข่ายทางสังคมของคณาจารย์คณะบริหารธุรกิจ ทั้งนี้เพื่อแสดงให้เห็นความสัมพันธ์ระหว่างคณาจารย์และเพิ่มคุณภาพของความสัมพันธ์ (Quality Relationship) ระหว่างคณาจารย์ภายในคณะบริหารธุรกิจให้มีการใช้เครือข่ายทางสังคมให้เกิดประสิทธิภาพมากยิ่งขึ้น โดยนำแนวทางการค้นหาค่า Network Centralization จาก [4] และข้อเสนอแนะของ [5] มาใช้เป็นแนวทางในการศึกษา

2. วรรณกรรมที่เกี่ยวข้อง

2.1 แนวคิดเกี่ยวกับเครือข่ายสังคม

เครือข่ายทางสังคม หมายถึง กลุ่มของสมาชิกเครือข่ายที่มีความเกี่ยวข้องเชื่อมโยงกัน โดยมีความสัมพันธ์และปฏิสัมพันธ์ทางสังคมในรูปแบบต่าง ๆ เช่น ความเป็นเพื่อน ญาติพี่น้อง ผู้ร่วมงาน การพูดคุย การให้คำแนะนำ การให้ความไว้วางใจ การแลกเปลี่ยนสารสนเทศ การติดต่อสื่อสารผ่านอีเมลและสื่อสังคมออนไลน์ การให้ความร่วมมือ ความเป็นผู้แต่งร่วม และการอ้างอิงผลงานวิจัย เป็นต้น [6]

โครงสร้างของเครือข่ายสามารถอธิบายได้ในรูปกราฟหรือตารางเมทริกซ์ โดยใช้สัญลักษณ์จุด แทนสมาชิกเครือข่าย และเส้น แทนความสัมพันธ์ โดยสมาชิกเครือข่ายและเส้นเชื่อมโยงอาจมีชื่อเรียกที่แตกต่างกันไปตามความนิยมของแต่ละสาขาวิชา ซึ่งล้วนมีความหมายเหมือนกัน เช่น สมาชิกเครือข่าย เรียกว่า actors, agents, egos, vertices, nodes, sites หรือ units เส้นเชื่อมโยงเรียกว่า relations, ties, alters, edges, arcs, links, connections, หรือ bonds เป็นต้น [7] ในส่วนของวิธีการในการวิเคราะห์เครือข่ายทางสังคมของคณาจารย์คณะบริหารธุรกิจนั้น ได้วัดค่าของความเป็นศูนย์กลาง (Centrality) ที่จะทำการค้นหาว่าใครหรือศูนย์รวมกิจกรรมใดมี “ความสำคัญ” ในเครือข่าย โดยตัวแสดงหรือศูนย์รวมกิจกรรมที่ว่าสำคัญนี้ จะวางตัวอยู่ในตำแหน่งอย่างมีนัยยะเชิงยุทธศาสตร์ของเครือข่าย [8] โดยจะทำการวัดค่าความเป็นศูนย์กลาง ได้แก่

1) Degree Centrality เป็นการค้นหาว่าสมาชิกใดบ้างที่เป็นจุดศูนย์กลางของการเชื่อมโยงซึ่งเป็นตำแหน่งที่มีอิทธิพลสูงสุดในเครือข่าย วัดได้จากจำนวนเส้นเชื่อมโยงทั้งหมดที่เชื่อมโยงมาจากสมาชิกเครือข่ายอื่น ๆ ทั้งที่อยู่ภายในกลุ่มเดียวกันและข้ามกลุ่ม [9] มีสมการดังนี้

$$d(i) = \sum_j m_{ij} \quad (1)$$

โดยที่ $m_{ij} = 1$ ถ้ามีการเชื่อมต่อระหว่างโหนดและ $m_{ij} = 0$ ถ้าไม่มีการเชื่อมต่อระหว่างกัน

2) Closeness Centrality เป็นการค้นหาว่าสมาชิกใดบ้าง ที่มีตำแหน่งเป็นจุดศูนย์กลางของเครือข่าย มีความใกล้ชิดกับสมาชิกอื่นและใช้ระยะทางที่สั้นที่สุดในการเข้าถึง ซึ่งวัดได้จาก

จำนวนเส้นเชื่อมโยงทั้งหมดที่ใช้ในการเดินทางจากสมาชิกหนึ่งไปยังอีกสมาชิกหนึ่ง โดยการลากผ่านสมาชิกอื่น ๆ ภายในเครือข่ายด้วยเส้นทางที่สั้นที่สุด มีสมการดังนี้

$$c(i) = \sum_j d_{ij} \quad (2)$$

โดย d_{ij} คือ จำนวนเส้นเชื่อมโยงในเส้นทางที่สั้นที่สุดจากสมาชิกหนึ่งไปยังอีกสมาชิกหนึ่ง

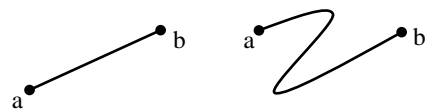
3) Betweenness Centrality เป็นการค้นหาว่าสมาชิกใดบ้างที่มีตำแหน่งเป็นสะพานเพื่อเชื่อมกลุ่มต่าง ๆ ที่อยู่ห่างกันให้เข้าหากัน หรือเป็นตัวกลางในการติดต่อเชื่อมโยงระหว่างสมาชิกอื่น ๆ [10] มีสมการดังนี้

$$b(i) = \sum_{j,k} \frac{g_{jik}}{g_{jk}} \quad (3)$$

โดย g_{jk} คือ จำนวนเส้นทางที่สั้นที่สุดจากโหนด j ไปยังโหนด k ($j, k \neq i$) และ g_{jik} คือ จำนวนเส้นทางที่สั้นที่สุดจากโหนด j ไปยังโหนด k ที่ต้องผ่านโหนด i

2.2 ทฤษฎีกราฟ (Graph Theory)

ทฤษฎีกราฟจะประกอบด้วยโหนดและลิงค์ [11] โดยกราฟเป็นแบบจำลองทางคณิตศาสตร์ ซึ่งใช้สำหรับจำลองปัญหาบางอย่างด้วยแผนภาพที่ประกอบด้วยโหนด และเส้นที่เชื่อมระหว่างจุด 2 จุด ในการเขียนแผนภาพของกราฟนั้น จะกำหนดตำแหน่งของจุดยอด ณ ตำแหน่งใดก็ได้ และจะลากเส้นเชื่อมของกราฟเป็นเส้นตรงหรือเส้นโค้งที่มีความยาวเป็นเท่าใดก็ได้ เช่น กำหนด เส้นเชื่อม ab ดังรูป (1) และอาจเขียนเส้นเชื่อม ab เป็นเส้นโค้ง ดังรูป (2) ได้ดังนี้



ภาพที่ 1: แสดงตัวอย่างการเขียนแผนภาพกราฟ [11]

2.3 โปรแกรม Pajek

การศึกษาโครงสร้างเครือข่ายทางสังคมของคณาจารย์คณะบริหารธุรกิจ ผู้วิจัยได้เลือกใช้โปรแกรม Pajek ในการจำลอง

ความสัมพันธ์ ทั้งนี้โปรแกรม Pajek สามารถช่วยในการวิเคราะห์และวาดภาพจำลองระบบเครือข่ายขนาดใหญ่ได้อย่างมีประสิทธิภาพ [12] โปรแกรม Pajek ถูกพัฒนาขึ้นมาด้วยภาษาปาสคาลโดย Vladimir Batagelj และ Andrej Mrvar โดยสามารถสร้างความสัมพันธ์เครือข่ายของการทำงานร่วมกันได้เป็นอย่างดี

โปรแกรม Pajek สามารถช่วยวิเคราะห์เครือข่ายที่มีขนาดใหญ่ได้ดีกว่าโปรแกรม MFinder, MAVisto, และ FANMOD ซึ่งโปรแกรมเหล่านี้ไม่มีความสามารถในการคำนวณค่าพารามิเตอร์ต่าง ๆ ของเครือข่ายขนาดใหญ่ได้ [13] จากงานวิจัยของ Andreas Harrer และ Alona Schmidt [14] ทำให้เกิดข้อค้นพบว่าการสร้างรูปแบบและกฎการเชื่อมโยงความสัมพันธ์หลาย ๆ เครือข่ายเข้าด้วยกันนั้น สามารถทำได้ในรูปแบบของ Blockmodelling ซึ่งโมเดลนี้ได้ใช้พื้นฐานการทำงานของโปรแกรม Pajek มาช่วยในการจำลองแผนภาพและวิเคราะห์ข้อมูล

3. ขอบเขตการดำเนินงาน

ผู้วิจัยได้นำทฤษฎีกราฟมากำหนดความสัมพันธ์ โดยกำหนดให้โหนด เป็นตัวแทนของบุคคลในโครงสร้าง ส่วนลิงค์ เป็นตัวแทนของความสัมพันธ์ระหว่างบุคคล ซึ่งในที่นี้คือความสัมพันธ์ระหว่างอาจารย์แต่ละท่าน โดยมีการเก็บข้อมูลความสัมพันธ์และสัมภาษณ์คณาจารย์เพื่อนำมาหาความสัมพันธ์ระหว่างบุคคล จากนั้นได้ประยุกต์ใช้โปรแกรม Pajek ในการจำลองเครือข่ายความสัมพันธ์ของคณาจารย์คณะบริหารธุรกิจและคำนวณค่าความเป็นจุดศูนย์กลางประกอบการพิจารณาความสำคัญของโหนดที่สำคัญที่สุดของเครือข่าย

4. ขั้นตอนการดำเนินงาน

การศึกษาครั้งนี้มีหน่วยการวิเคราะห์ระดับกลุ่ม คือเครือข่ายทางสังคมของคณาจารย์คณะบริหารธุรกิจ ซึ่งมีทั้งหมด 137 คน ประกอบด้วย สาขาวิชา ได้แก่ สาขาวิชา 9 บริหารธุรกิจดุสิตบัณฑิต สาขาวิชาบริหารธุรกิจมหบัณฑิต สาขาวิชาคอมพิวเตอร์ธุรกิจ สาขาวิชาการตลาด สาขาวิชาการจัดการอุตสาหกรรม สาขาวิชาการจัดการ สาขาวิชาการจัดการ

การกีฬา สาขาวิชาการจัดการ โลจิสติกส์ และสาขาวิชาธุรกิจ อสังหาริมทรัพย์

ผู้วิจัยเก็บรวบรวมข้อมูล โดยใช้การสัมภาษณ์คณาจารย์คณะบริหารธุรกิจจำนวน 137 คน สามารถแสดงขั้นตอนการดำเนินงาน ได้ดังนี้

ขั้นตอนที่ 1 ทำการรวบรวมข้อมูลจำนวนคณาจารย์ในสาขาต่าง ๆ ของคณะบริหารธุรกิจ มหาวิทยาลัยกรุงเทพธนบุรี

ขั้นตอนที่ 2 ดำเนินการสัมภาษณ์คณาจารย์คณะบริหารธุรกิจเพื่อค้นหาความสัมพันธ์กันของคณาจารย์

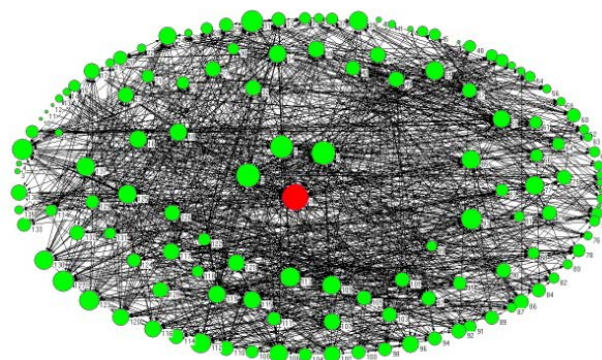
ขั้นตอนที่ 3 นำความสัมพันธ์คณาจารย์มาจัดทำเป็น Network Data File เพื่อเตรียมเข้าสู่โปรแกรม Pajek

ขั้นตอนที่ 4 คำนวณค่าความเป็นจุดศูนย์กลาง ได้แก่ Degree Centrality, Closeness Centrality และ Betweenness Centrality ด้วยโปรแกรม Pajek

ขั้นตอนที่ 5 การนำเสนอผลการดำเนินงานพร้อมแสดงภาพจำลองความสัมพันธ์ของคณาจารย์คณะบริหารธุรกิจ

5. ผลการดำเนินงาน

จากการศึกษาครั้งนี้ผู้วิจัยได้จำลองแผนภาพความสัมพันธ์ของคณาจารย์คณะบริหารธุรกิจ โดยใช้ทฤษฎีกราฟมาช่วยในการกำหนดความสัมพันธ์ และใช้โปรแกรม Pajek ในการสร้างแผนภาพความสัมพันธ์ ซึ่งสามารถแสดงได้ดังภาพที่ 3



ภาพที่ 2: ความสัมพันธ์ของคณาจารย์

จากภาพที่ 2 แสดงความสัมพันธ์ของคณาจารย์คณะบริหารธุรกิจ โดยมีจำนวน 137 โหนดตามจำนวนคณาจารย์ และในแต่ละโหนดจะมีเส้นหรือลิงค์เพื่อแทนความสัมพันธ์ของคณาจารย์ว่าใครสามารถกระจายข้อมูล

ข่าวสารหรือมีความสัมพันธ์กับใครบ้าง ในลักษณะทิศทางเดียวกันทั้งหมด

ในการค้นหาว่าอาจารย์ท่านใด มีความสำคัญและเป็นตัวกลางในการติดต่อสื่อสารหรือการแลกเปลี่ยนข้อมูล จำเป็นต้องพิจารณาค่าความศูนย์กลาง (Centrality) โดยแสดงผลการวิเคราะห์ข้อมูลได้ดังนี้

ตารางที่ 1: ค่า Degree Centrality

1. All Degree Centrality of N1 (137)			
The lowest value: 2			
The highest value:76			
Vertex	Id	Value	Rank
1	1BM	76	1
2	2BM	64	2
5	5BM	62	3
7	7BM	61	4

จากตารางที่ 1 พบว่า โหนดที่ 1 หรืออาจารย์ 1BM คือศูนย์กลางรวมความสัมพันธ์ของคณาจารย์คนอื่น ๆ ซึ่งเป็นบุคคลที่มีความสำคัญในลำดับต้น ๆ เมื่อต้องการกระจายข้อมูลข่าวสาร ควรพิจารณาอาจารย์ 1BM เป็นอันดับแรกเนื่องจากมีจำนวนของความสามารถในการกระจายข่าวสารถึงคนอื่น ๆ มากที่สุดคือ 76 รองลงมาคืออาจารย์ 2BM และ 5BM ตามลำดับ

ตารางที่ 2: ค่า Closeness Centrality

2. All Closeness Centrality of N1 (137)			
The lowest value: 0.377			
The highest value:0.60 2			
Vertex	Id	Value	Rank
1	1BM	0.602	1
2	2BM	0.586	2
5	5BM	0.584	3
6	6BM	0.569	4

จากตารางที่ 2 พบว่า คณาจารย์ที่สามารถส่งผ่านข้อมูลไปยังคณาจารย์ท่านอื่น ๆ ได้อย่างรวดเร็วที่สุดคือ อาจารย์ 1BM โดยมีค่า Closeness Centrality มากที่สุดเป็น 0.602 ซึ่งเมื่อโหนดใดมีค่า Closeness Centrality สูงย่อมหมายถึงความสามารถในการติดต่อกับคณาจารย์ท่านอื่น ๆ นั้นเป็นไปได้อย่างรวดเร็ว รองลงมาคืออาจารย์ 2BM และ 5BM ตามลำดับ

ตารางที่ 3: ค่า Betweenness Centrality

3. Betweenness Centrality of N1 (137)			
The lowest value: 0			
The highest value:0.163			
Vertex	Id	Value	Rank
1	1BM	0.163	1
2	2BM	0.086	2
7	7BM	0.067	3
5	5BL	0.057	4

จากตารางที่ 3 พบว่า อาจารย์ 1BM มีค่า Betweenness สูงที่สุด โดยมีค่าเป็น 0.163 ดังนั้นอาจารย์ 1BM จึงเป็นผู้ที่มีความสำคัญหรือเป็นศูนย์กลางรวมของการกระจายข่าวสาร และสามารถทำหน้าที่คล้ายกับตัวกลางในการจัดสรรอำนาจ แม้ว่าจะไม่มีการเชื่อมต่อโดยตรงก็ตาม

6. บทสรุป

เครือข่ายทางสังคม (Social Network) สามารถแสดงถึงรูปแบบความสัมพันธ์ระหว่างคณาจารย์ภายในคณะบริหารธุรกิจได้อย่างชัดเจน โดยใช้ทฤษฎีกราฟในการกำหนดโหนดหรือจุดแต่ละจุด ซึ่งในที่นี้จะแทนคณาจารย์ และใช้ลิงค์หรือเส้น เชื่อมโยง เพื่อแสดงถึงความสัมพันธ์จากอาจารย์ท่านหนึ่งไปยังอีกท่านหนึ่ง

การนำโปรแกรม Pajek มาใช้ในการวิเคราะห์เครือข่ายทางสังคมของคณาจารย์คณะบริหารธุรกิจในครั้งนี้ เป็นการจำลองความสัมพันธ์โดยอยู่บนพื้นฐานของทฤษฎีกราฟ และการศึกษาค้นคว้าครั้งนี้ ได้วิเคราะห์หาค่า Degree Centrality, Closeness Centrality และ Betweenness Centrality จากการหาค่าดังกล่าว

ทำให้พบว่า อาจารย์ที่สำคัญที่สุดภายในเครือข่าย ได้แก่ อาจารย์ IBM ทั้งนี้ เมื่อต้องการส่งข้อมูล ข่าวสารต่าง ๆ ไปยังคณาจารย์ท่านอื่น ๆ ควรพิจารณาถึงอาจารย์ IBM เป็นหลัก เพื่อให้เป็นตัวกลางในการกระจายข้อมูลข่าวสาร ซึ่งโอกาสที่คณาจารย์ในคณะบริหารธุรกิจ จะได้รับข้อมูลข่าวสารอย่างรวดเร็ว นั้น จะเป็นไปได้มากกว่าการกระจายข้อมูลข่าวสารจากอาจารย์ที่มีค่าความเป็นศูนย์กลาง (Centrality) ต่ำ

เอกสารอ้างอิง

- [1] รุจเรขา วิทยาวุฒิกุล และน้ำทิพย์ วิกาวิน, “การวิเคราะห์เครือข่ายทางสังคมกับงานวิจัยทางสารสนเทศศาสตร์,” *Journal of Library and Information Science*, ปีที่ 5 ฉบับที่ 2 กรกฎาคม-ธันวาคม 2555 หน้า 126-140.
- [2] M. E. J., Newman, “The Structure and Function of Complex Network,” *SIAM Review* 45, pp. 167-256, 2003.
- [3] ศิริลักษณ์ โรจนกิจอำนวย, “องค์ประกอบคุณภาพความสัมพันธ์ของเครือข่ายสังคม,” *วารสารบริหารธุรกิจ*, ปีที่ 35 ฉบับที่ 133 มกราคม-มีนาคม 2555 หน้า 9-18.
- [4] วิทยา เรืองฤทธิ์, “การวิเคราะห์เครือข่ายทางสังคมของกระแสดูหนังเดินทางไปทำงานในภูมิภาคเมืองโคราช,” *Veridian E-Journal มหาวิทยาลัยศิลปากร* ปีที่ 5 ฉบับที่ 3 กันยายน-ธันวาคม 2555 หน้า 254-266.
- [5] ทัดชนันท์ พุ่มนุช, “การศึกษาพฤติกรรมการใช้เครือข่ายสังคม (Social Network) เพื่อพัฒนาในการปฏิบัติงานของบุคลากรทางการศึกษาในสำนักงานเขตพื้นที่การศึกษาประถมศึกษานครปฐม,” *Veridian E-Journal*, มหาวิทยาลัยศิลปากร ปีที่ 5 ฉบับที่ 1 มกราคม-เมษายน 2555 หน้า 523-540
- [6] A. Marin and B. Wellman, “Social Network Analysis: An Introduction,” *The SAGE Handbook of Social Network Analysis*, Edited by J. Scott and P. J. Carrington, pp. 11-25, 2011.
- [7] B. Hoppe and C. Reinelt, “Social network analysis and the evaluation of leadership networks,” *The Leadership Quarterly*, vol. 21, no. 4, pp. 600-619, 2010.
- [8] S. Wasserman and K. Faust, “Social Network Analysis: Methods and Applications,” *The Press Syndicate of the University of Cambridge*, 1999.
- [9] R. A. Hanneman and M. Riddle, “Introduction to social network methods,” *University of California*, 2005.
- [10] M. M. Durland, K. A. Fredericks, and American Evaluation Association, “Social network analysis in program evaluation,” *Jossey-Bass*, Calif, San Francisco, 2006.
- [11] W. de Nooy, A. Mrvar, and V. Batagelj, “Exploratory social network analysis with Pajek,” *Cambridge University Press*, New York, 2005.
- [12] V. Batagelj and A. Mrvar, “Pajek — Analysis and Visualization of Large Networks in Graph Drawing Software,” *Springer Berlin Heidelberg*, pp. 77-103, 2004.
- [13] T. Milenkovic, J. Lai, and N. Pržulj, “GraphCrunch: A tool for large network analyses,” *BMC Bioinformatics*, vol. 9, no. 1, p. 70, Jan. 2008.
- [14] A. Harrer and A. Schmidt, “Blockmodelling and role analysis in multi-relational networks,” *Springer-Verlag Wien.*, vol. 3, no. 3, pp. 701-719, Sep. 2013.

TCSEdPM: โดเมนออนโทโลยีสำหรับการบริหารงานบุคคล ข้าราชการครูและบุคลากรทางการศึกษา

TCSEdPM: A Domain Ontology for the Teacher Civil Service and Educational Personnel Management

ดวงพร เจียมอัมพร (Duangporn Jeamumporn)¹ สุพจน์ นิตยสุวรรณ (Supot Nitsuwat)² และมาลีรัตน์ โสदानิล (Maleerat Sodanil)¹

¹คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ กรุงเทพฯ

²ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ประยุกต์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ กรุงเทพฯ

d_tum@hotmail.com, sns@kmutnb.ac.th, msn@kmutnb.ac.th

บทคัดย่อ

บทความนี้นำเสนอการพัฒนา TCSEdPM โดเมนออนโทโลยีสำหรับการบริหารงานบุคคลข้าราชการครูและบุคลากรทางการศึกษา เนื่องจากการบริหารงานบุคคลเป็นหัวใจของการบริหารงาน ข้อมูลและองค์ความรู้ต่างๆ ที่เกี่ยวข้องมีอยู่เป็นจำนวนมาก หลากหลายรูปแบบ แต่ขาดการจัดการความรู้ที่ดี ทำให้การสืบค้นเพื่อนำองค์ความรู้มาใช้ให้เกิดประโยชน์มีความยุ่งยาก ออนโทโลยี เป็นแนวคิดในการพัฒนาและนำฐานความรู้มาช่วยอำนวยความสะดวกแก่ผู้ใช้งาน และทำให้ระบบงานมีความชาญฉลาดและมีประสิทธิภาพยิ่งขึ้น การพัฒนา TCSEdPM ออนโทโลยี แบ่งเป็น 3 กระบวนการ คือ 1) การถอดองค์ความรู้ 2) การออกแบบและพัฒนาออนโทโลยีด้วยโปรแกรม Hozo-Ontology Editor และเชื่อมต่อระหว่างฐานข้อมูลกับฐานข้อมูลออนโทโลยี และ 3) การพัฒนาระบบสืบค้นข้อมูลเชิงความหมาย และระบบให้คำแนะนำซึ่งเป็นขั้นตอนต่อไปของงานวิจัย ซึ่งผลการพัฒนาได้ TCSEdPM โดเมนออนโทโลยีที่แบ่งแนวคิดที่มีความเกี่ยวข้องออกเป็น 11 แนวคิด

คำสำคัญ: TCSEdPM โดเมนออนโทโลยี การบริหารงานบุคคล

Abstract

This paper reports on the development of TCSEdPM a domain ontology for the Teacher Civil Service and Educational Personnel Management. Because Personnel Management is the heart of administration, but they are lack of good management. Hence, the retrieval of

knowledge are difficult to do. Ontology is a concept of knowledge based development to facilitate the user more conveniently and it's make the systems smarter and more powerful. TCSEdPM development process can be divided into three processes with 1) knowledge capture 2) design ontology and development with Hozo-Ontology Editor and database-ontology mapping, and last process 3) development application, a semantic ontology search and advisory system which be the next step of the research. Which the results of development is TCSEdPM Domain Ontology, dividing related concepts in to eleven.

Keyword: TCSEdPM, domain ontology, personnel management.

1. บทนำ

การบริหารงานบุคคลเป็นหัวใจของการบริหารงานในทุกหน่วยงาน ปัญหาของงานการบริหารงานบุคคลที่เกิดขึ้นแล้วในอดีตสามารถเกิดขึ้นซ้ำๆ ได้อีกในอนาคต เช่น ปัญหาเรื่องการย้ายข้าราชการครูและบุคลากรทางการศึกษา ปัญหาเรื่องการรับโอนข้าราชการพลเรือนสามัญ ปัญหาการใช้บัญชีผู้สอบแข่งขันได้ เป็นต้น และพบว่า วิธีการแก้ปัญหากรณีในอดีตสามารถนำกลับมาใช้แก้ปัญหที่เกิดขึ้นใหม่ได้อย่างมีประสิทธิภาพ วิธีการแก้ปัญหาโดยการพัฒนาโดเมนออนโทโลยี (Domain Ontology) เป็นศาสตร์ที่ใช้ในการพัฒนาฐานความรู้ที่มีลักษณะเชิงเนื้อหา (Content Base) [1], [2], [3] นำประโยชน์ของออนโทโลยีมาใช้เพื่ออธิบายและตรวจสอบความหมายที่เหมือนกัน และเมื่อนำมาประยุกต์ใช้ร่วมกับเทคนิคการให้เหตุผลด้วยกรณีศึกษา (Case-Based Reasoning: CBR)

เป็นวิธีการค้นหากรณีปัญหาเดิมจากฐานองค์ความรู้ที่มีลักษณะใกล้เคียงกับกรณีปัญหาใหม่มากที่สุด นำเอาวิธีการแก้ปัญหาที่ได้จากกรณีเดิมที่ค้นหาได้ ไปประยุกต์ใช้เพื่อแก้ปัญหาใหม่ สำหรับการพัฒนาระบบให้คำแนะนำ (Advisory System) โดยฐานความรู้ออนโทโลยีจะช่วยให้ส่วนประมวลผลเชิงอนุมาน (Inference Engine) สำหรับสร้างผลลัพธ์การแนะนำข้อมูล ทำงานได้อย่างมีความแม่นยำและมีประสิทธิภาพมากขึ้น [3] สามารถนำเสนอคำตอบที่ถูกต้องให้กับผู้ใช้งานมากที่สุด

ผู้วิจัยจึงมีแนวคิดริเริ่มที่จะพัฒนาโดเมนออนโทโลยีขึ้นเพื่อจัดเก็บข้อมูลที่มีการนำเสนอเนื้อหาแบบมีโครงสร้าง มีขอบเขตเนื้อหาเฉพาะทางของงานการบริหารงานบุคคลของข้าราชการครูและบุคลากรทางการศึกษา ด้วยการสร้างฐานองค์ความรู้และกรณีศึกษาเพื่อใช้ในการประมวลคำถาม (Problem) และค้นคืนคำตอบ (Solution) สำหรับระบบให้คำแนะนำ (Advisory System) งานการบริหารงานบุคคล ข้าราชการครูและบุคลากรทางการศึกษาของกระทรวงศึกษาธิการ โดยใช้งานร่วมกับเทคนิคการให้เหตุผลด้วยกรณีศึกษา เพื่อให้ผู้ใช้งานสามารถสืบค้นคำตอบที่ต้องการผ่านเว็บไซต์และจะต้องได้รับคำแนะนำที่มีความเหมาะสมและถูกต้องมากที่สุด และจะทำการศึกษาการยอมรับเทคโนโลยีใหม่เข้าไปใช้ในองค์กร (Technology Acceptance) โดยเก็บข้อมูลจากผู้ใช้งานให้คำแนะนำในการปฏิบัติงานในโอกาสต่อไป

2. ทฤษฎีที่เกี่ยวข้อง

2.1 การบริหารงานบุคคล

กระบวนการดำเนินงานเกี่ยวกับบุคคลในหน่วยงาน เพื่อให้ได้มาซึ่งบุคคลที่มีความรู้ความสามารถเหมาะสมกับความต้องการของหน่วยงาน การจัดการกับบุคคล หรือการบริหารงานบุคคลเป็นหัวใจของการบริหารงาน หากการบริหารงานบกพร่อง ประสิทธิภาพของงานก็จะบกพร่อง แต่ถ้าได้รับความสำเร็จ ผลงานก็จะมีประสิทธิภาพด้วย ดังนั้น การบรรลุเป้าหมายของงานการบริหารงานบุคคลอย่างมีประสิทธิภาพจึงเป็นประเด็นปัญหาที่สำคัญ คือทำอย่างไรจึงจะทำให้ได้บุคลากรที่ดี สามารถบำรุงรักษาให้บุคลากรอยู่กับหน่วยงานนานๆ และบุคลากรเหล่านั้นได้รับการส่งเสริมให้มีความเจริญก้าวหน้า สามารถบรรลุเป้าหมายของการบริหารงานบุคคลได้อย่างมีประสิทธิภาพ และสามารถประหยัคงบประมาณของรัฐ สำหรับกระทรวงศึกษาธิการ

เป็นองค์กรที่มีขนาดใหญ่ และมีความหนาแน่นทรัพยากรบุคคลเป็นจำนวนมาก แต่การบริหารงานบุคคลของข้าราชการครูและบุคลากรทางการศึกษา ซึ่งดำเนินการโดยสำนักงาน ก.ค.ศ. มักประสบปัญหาต่างๆมากมาย [4] เนื่องจากผู้ที่มีหน้าที่ตอบคำถามมีความเข้าใจเกี่ยวกับงานที่คลาดเคลื่อน ซึ่งทำให้เกิดผลเสียต่อข้าราชการครูและบุคลากรทางการศึกษาเอง เจ้าหน้าที่ผู้ปฏิบัติงานและหน่วยงานต่างๆ ที่เกี่ยวข้อง

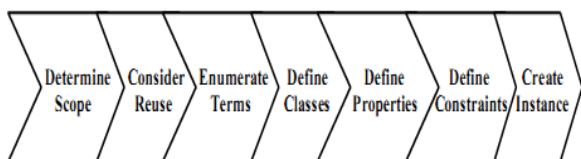
2.2 โดเมนออนโทโลยี

วิธีการบรรยายความรู้อย่างมีขอบเขตเป็นการสร้างโครงสร้างฐานความรู้ใช้ในการอธิบายความหมายของสิ่งต่างๆ และจัดหมวดหมู่เอกสารของข้อมูลได้ในขอบเขตโดยกฏนิยามเพื่อกำหนดโครงสร้าง ความสนใจหนึ่งๆ และความสัมพันธ์ของสิ่งที่สนใจ (Domain) [1], [2], [3] ออนโทโลยีสามารถประยุกต์ใช้กับงานหลาย ๆ ด้าน เช่น เว็บเชิงความหมาย การจัดการองค์ความรู้ พาณิชย์อิเล็กทรอนิกส์ และการค้นคืนสารสนเทศ

ข้อดีของออนโทโลยี [1], [2], [3] คือ สามารถแก้ปัญหาของฐานข้อมูลเชิงสัมพันธ์ (Relational Database) ที่ไม่สามารถค้นหาข้อมูลที่มีความสัมพันธ์กัน ออนโทโลยีสามารถนำมาใช้ในการจัดการกับการค้นหาคำศัพท์ที่เกี่ยวข้องเพราะใช้แบบจำลองแนวคิดสำหรับการค้นหาคำศัพท์ที่เกี่ยวข้อง (Relevant Terms) กับเอนทิตี (Entities) และความสัมพันธ์ (Relationship) ความสัมพันธ์ที่สัมพันธ์กัน (Related-to-Relationship) ในการบูรณาการฐานข้อมูล และการเชื่อมโยงคำศัพท์ของแบบจำลองฐานข้อมูล (Database Models) เพื่อตรวจสอบความหมายที่เหมือนกัน ทำให้ระบบมีประสิทธิภาพในการสืบค้นเพิ่มมากขึ้น สามารถนำเสนอคำตอบที่ถูกต้องและแม่นยำให้กับผู้ใช้งานมากที่สุด

ออนโทโลยีเป็นลักษณะภาษาที่นำมาใช้บรรยายโครงสร้างและความสัมพันธ์ของระบบ [5], [6] ผ่านโหนดแบบลำดับชั้น (Hierarchies) มีการกำหนดภาษามาตรฐานที่ใช้จำลองและออกแบบโครงสร้างของเอกสารเอ็กซ์เอ็มแอล (XML) โดยใช้นิยามแนวคิดให้อยู่ในรูปของกฎ (Rule) คลาส (Class) ความสัมพันธ์ระหว่างคลาส (Relationship) และคุณสมบัติของคลาส (Properties) ภาษาที่ใช้แทนข้อมูลเชิงความหมาย เช่น ภาษา XML, ภาษา RDF, ภาษา RDFS และภาษา OWL

กระบวนการพัฒนาออนโทโลยี สามารถแบ่งออกได้เป็น 7 ขั้นตอน [1] ดังแสดงในภาพที่ 1



ภาพที่ 1: กระบวนการพัฒนาออนโทโลยี

จากภาพที่ 1 กระบวนการพัฒนาออนโทโลยี เริ่มต้นด้วยการระบุขอบเขตและวัตถุประสงค์ของการพัฒนาออนโทโลยี (Determine Scope) หลังจากนั้นเป็นการพิจารณานาออนโทโลยีที่มีอยู่กลับมาใช้ใหม่ (Consider Reuse) การพิจารณาและแจกแจงรายละเอียดของเทอมหรือคำทั้งหมด (Enumerate Terms) การกำหนดคลาสและลำดับชั้นที่อยู่ในโดเมน (Define Classes) การกำหนดคุณสมบัติของคลาสและสล็อต (Define Properties) การกำหนดเงื่อนไขให้กับคุณสมบัติหรือกำหนดฟาเซ็ทให้กับสล็อต (Define Constrains) ตามลำดับ โดยมีขั้นตอนสุดท้ายคือ การสร้างตัวอย่างข้อมูลในแต่ละคลาส (Create Instance) ปัจจุบันมีเครื่องมือสำหรับพัฒนาออนโทโลยีมากมาย แต่ละเครื่องมือสนับสนุนการทำงานที่แตกต่างกัน โปรแกรมช่วยสร้างออนโทโลยี (Ontology Editor) เช่น Hozo-Ontology Editor [7] และ Protégé [8] เป็นต้น

สำหรับงานวิจัยที่เกี่ยวข้องกับออนโทโลยีมีมากมายเช่น การพัฒนาออนโทโลยีเพื่อประยุกต์ใช้กับฐานข้อมูลเชิงสัมพันธ์ที่หลากหลาย [9] โดยกำหนดคำศัพท์ คุณลักษณะของความสัมพันธ์ระหว่างคำศัพท์และความหมายที่ใช้ร่วมกัน และให้ข้อเสนอแนะว่า การพัฒนาออนโทโลยีเป็นกระบวนการที่จำเป็นต้องทำซ้ำ และยังมีการพัฒนาออนโทโลยีเพื่อบูรณาการฐานข้อมูลด้านภูมิศาสตร์ [10] เพื่อแก้ปัญหาการใช้และจัดการข้อมูลจากแหล่งข้อมูลที่มีความหลากหลายและแหล่งข้อมูลที่ไม่สัมพันธ์กันด้วยวิธีการบูรณาการระบบและบูรณาการข้อมูล เป็นวิธีการที่ง่าย แต่ใช้การลงทุนที่สูง ทำการรวมคำศัพท์ที่สมบูรณ์และการบรรยายความหมายที่สื่อให้เกิดความเข้าใจ มีประสิทธิภาพสอดคล้องและตรงตามความต้องการ

มีนักวิจัยหลายท่านซึ่งได้นำออนโทโลยีมาประยุกต์ใช้กับระบบงานหลากหลาย เช่น ระบบการดูแลสุขภาพทางด้านโภชนาการอาหารได้หวัน [5] โดยใช้ 3 ตัวแทน ได้แก่ ความรู้ส่วนบุคคล ร่วมกับ การหาข้อสรุปผลลัพธ์แบบพีชชี และเว็บ/เชิงความหมาย สำหรับประเมินสุขภาพของอาหารรับประทานเพื่อการควบคุมน้ำหนัก ผลการทดลองพบว่า ระบบมีประสิทธิภาพสามารถใช้เป็นข้อมูลอ้างอิงเพื่อส่งเสริมการมีสุขภาพที่ดีได้ รวมทั้งระบบให้คำแนะนำการบริโภคอาหารตามโภชนาการเฉพาะบุคคล เพื่อเก็บข้อมูลองค์ความรู้ของระบบแนะนำให้มีโครงสร้างที่สัมพันธ์กันระหว่างเนื้อหาและความรู้ [7] โดยใช้ฐานความรู้ออนโทโลยีร่วมกับฐานกุนำความรู้เฉพาะทางของผู้เชี่ยวชาญด้านโภชนาการ สำหรับให้คำแนะนำการบริโภคอาหารสำหรับผู้ป่วยเป็นโรคเบาหวาน เพื่อให้ผู้ป่วยสามารถรับประทานอาหารที่ถูกต้องตามหลักโภชนาการผ่านทางเว็บเบราว์เซอร์ได้ นอกจากนี้ยังมีงานวิจัยเกี่ยวกับออนโทโลยีกับการพัฒนาเว็บเชิงความหมาย เช่น การพัฒนาโดเมนออนโทโลยีเกี่ยวกับความสนใจของผู้ใช้บริการสำหรับการค้นหาบนเว็บเชิงความหมายชื่อว่า OntoSearch [11] พบว่า สามารถให้บริการข้อมูลและเอกสารตามความสนใจของแต่ละบุคคลตามมาตรฐานของห้องสมุดดิจิทัล ACM ได้อย่างมีประสิทธิภาพ

จากการทบทวนงานวิจัยที่เกี่ยวข้องกับออนโทโลยีข้างต้น ผู้วิจัยสามารถที่จะนำแนวคิดที่มีประโยชน์ดังกล่าวมาแล้วมาประยุกต์ใช้สำหรับการแก้ปัญหาการค้นหาข้อมูลจากฐานข้อมูลเชิงสัมพันธ์ด้วยการนำออนโทโลยีมาใช้ เพื่อเพิ่มประสิทธิภาพให้กับระบบงาน

2.3 Hozo-Ontology Editor

โปรแกรม Hozo-Ontology Editor [7], [12] เป็นโปรแกรมที่พัฒนาขึ้นโดยมหาวิทยาลัยโอซากา (Osaka University) ประเทศญี่ปุ่น เป็นโปรแกรมเครื่องมือสนับสนุนการพัฒนาออนโทโลยี (Ontology Editor) ซึ่งได้รับความนิยมในปัจจุบันเนื่องจากเป็นเครื่องมือสำหรับจัดเก็บองค์ความรู้ในรูปแบบของออนโทโลยีได้สะดวกและง่ายยิ่งขึ้น

3. วิธีการดำเนินงาน

งานวิจัยนี้ประกอบด้วย 2 ส่วน คือ ส่วนการศึกษาและวิเคราะห์ปัญหา และส่วนการออกแบบและพัฒนาออนโทโลยี

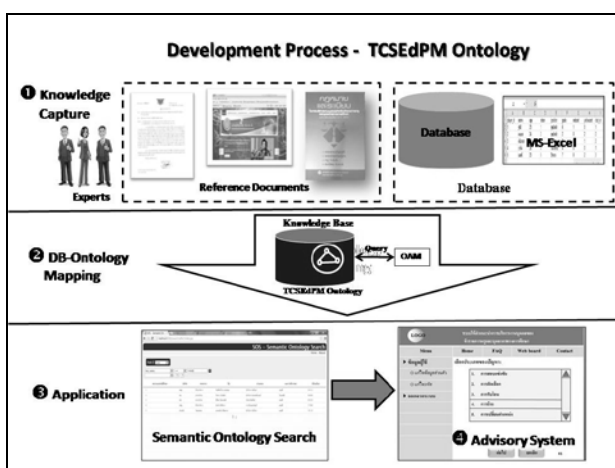
3.1 การศึกษาและวิเคราะห์ปัญหา

จากการศึกษาการดำเนินงานการบริหารงานบุคคลของ สنج.ศ.ค.ก. [4] พบว่า กรณีปัญหาและคำถามมีจำนวนมากและเกิดขึ้นซ้ำ ๆ ข้อมูลการแก้ปัญหาอยู่ในรูปแบบของเอกสาร และผู้เชี่ยวชาญ เมื่อมีปัญหเกิดขึ้นเจ้าหน้าที่เกี่ยวข้องจะเป็น ผู้ให้คำตอบหรือแนวทางแก้ไขอาจมีความเข้าใจที่คลาดเคลื่อน ทำให้ไม่สามารถตอบสนองความต้องการของผู้รับบริการ

ผู้วิจัยจึงได้นำประโยชน์ของฐานความรู้ออนโทโลยี ระบบมาให้คำแนะนำมาประยุกต์ใช้กับการพัฒนา เพื่อปรับปรุง กระบวนการแก้ไขปัญหาที่เกิดขึ้นกับงานการบริหารบุคคล ดังที่กล่าวมาแล้ว

3.2 การออกแบบและพัฒนาออนโทโลยี

เมื่อทำการศึกษาและวิเคราะห์ปัญหาแล้ว ผู้วิจัยจึงได้ออกแบบโดยนำเสนอ TCSEdPM เป็นโดเมนออนโทโลยี ที่รวบรวมความรู้เฉพาะทางด้านการบริหารงานบุคคล ข้าราชการครูและบุคลากรทางการศึกษา กรณีศึกษา สำนักงาน ก.ค.ศ. กระทรวงศึกษาธิการ [4] เนื่องจากออนโทโลยีการบริหารงานบุคคลข้าราชการครูและบุคลากรทางการศึกษา ยังไม่เคยมีผู้ใดจัดทำมาก่อน ผู้วิจัยจึงได้ดำเนินการออกแบบและพัฒนา TCSEdPM โดเมนออนโทโลยีสำหรับการบริหารงานบุคคล โดยแบ่งเป็น 3 กระบวนการ ดังแสดงในภาพที่ 2



ภาพที่ 2: กระบวนการออกแบบและพัฒนา TCSEdPM ออนโทโลยี

จากภาพที่ 2 กระบวนการออกแบบและพัฒนา TCSEdPM ออนโทโลยี แบ่งออกเป็น 3 กระบวนการ คือ กระบวนการแรก เป็นกระบวนการถอดองค์ความรู้ (Knowledge Capture) การรวบรวมความรู้ประสบการณ์ในการปฏิบัติงานด้าน

การบริหารงานบุคคลข้าราชการครูและบุคลากรทางการศึกษา จากผู้เชี่ยวชาญหรือเจ้าขององค์ความรู้ (Expert) เอกสาร (Reference Documents) และข้อมูลที่อยู่บนเว็บไซต์ของหน่วยงาน รวมถึงฐานข้อมูลทั้งที่เป็นแบบอิเล็กทรอนิกส์ (Database) และฐานข้อมูลแบบตารางที่สร้างด้วยโปรแกรม MS-Excel หลังจากนั้นนำองค์ความรู้ที่ได้มาสรุปในลักษณะของผังความคิด (Mind Map) และสร้างเป็นแบบจำลอง Conceptual Model โดยนำข้อมูลที่ได้อาวิเคราะห์และสร้างโครงสร้างข้อมูลในรูปแบบของความสัมพันธ์ของชั้นข้อมูล แบ่งออกเป็น ชั้นข้อมูลหลักเชื่อมต่อกับชั้นข้อมูลรองและชั้นข้อมูลย่อย

หลังจากนั้นทำการตรวจสอบความถูกต้องขององค์ความรู้ ด้วยวิธีการประเมินความถูกต้อง ความเหมาะสม และความสมบูรณ์ขององค์ความรู้ ด้วยการประเมินเนื้อหา (Content) โดยผู้เชี่ยวชาญของสำนักงานคณะกรรมการข้าราชการครูและบุคลากรทางการศึกษา กระทรวงศึกษาธิการจำนวน 3 คน เครื่องมือที่ใช้ คือ แบบประเมินเนื้อหาการออกแบบออนโทโลยี งานการบริหารงานบุคคลข้าราชการครูและบุคลากรทางการศึกษา

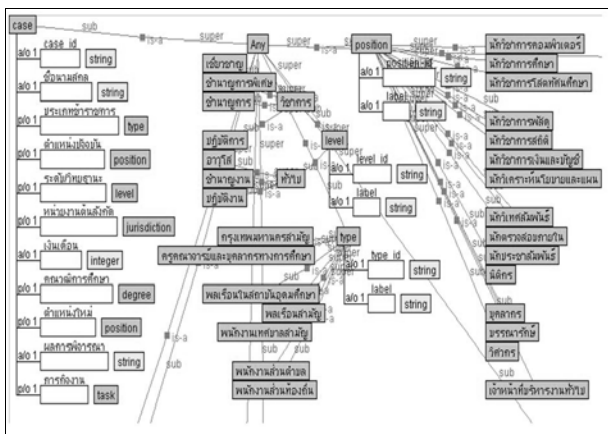
โดยมีรายละเอียดของการพิจารณาและข้อเสนอแนะของผู้เชี่ยวชาญตามตารางสรุปผลการพิจารณาของผู้เชี่ยวชาญด้านเนื้อหา ดังแสดงในตาราง 1 ดังนี้

ตารางที่ 1: สรุปผลการพิจารณาของผู้เชี่ยวชาญด้านเนื้อหา

หัวข้อ (Class)	ผลการพิจารณา
1. วิธีการบรรจุและแต่งตั้ง	เหมาะสม 2 คน/ปรับปรุง 1 คน
2. บุคลากร	เหมาะสม 2 คน/ปรับปรุง 1 คน
3. คุณสมบัติ	เหมาะสม 2 คน/ปรับปรุง 1 คน
4. ตำแหน่ง	ปรับปรุง 3 คน
5. สังกัด	เหมาะสม 2 คน/ปรับปรุง 1 คน
6. ชั้นเงินเดือน	เหมาะสม 3 คน
7. คุณวุฒิการศึกษา	เหมาะสม 3 คน
8. สาขาวิชาเอก	เหมาะสม 3 คน
9. จำนวนหน่วยกิต	เหมาะสม 3 คน
10. ระยะเวลา	ปรับปรุง 3 คน
11. เอกสารอ้างอิง	เหมาะสม 3 คน

จากตารางที่ 1 นำผลการประเมิน คำแนะนำและ ข้อเสนอแนะของผู้เชี่ยวชาญด้านเนื้อหา เช่น ควรปรับปรุงการนิยามความหมาย การเพิ่มหัวข้อย่อย และรายการข้อมูลที่เป็นไปได้/ตัวอย่างข้อมูลไปปรับปรุงและแก้ไขการออกแบบและสร้าง TCSEdPM ออนโทโลยีให้มีความเหมาะสมมากขึ้น

กระบวนการที่ 2 กระบวนการออกแบบและพัฒนา TCSEdPM ออนไลน์ และการเชื่อมต่อระหว่างฐานข้อมูลและฐานข้อมูลออนไลน์ (DB-Ontology Mapping) การพัฒนา TCSEdPM ออนไลน์ โดยนำโครงสร้างของแบบจำลองมาพัฒนาให้อยู่ในรูปแบบของภาษา OWL โดยใช้โปรแกรม Hozo-Ontology Editor Version 5.2.36 [7], [12] ซึ่งเป็นโปรแกรมสำหรับถ่ายถอดและจัดเก็บองค์ความรู้ในรูปแบบออนไลน์ ดังแสดงในภาพที่ 3



ภาพที่ 3: ตัวอย่างการออกแบบ TCSEdPM ออนไลน์

จากภาพที่ 3 แสดงตัวอย่างการออกแบบ TCSEdPM ออนไลน์ ด้วยโปรแกรม Hozo-Ontology Editor

```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns="http://www.hozo.jp/owl/TEdPM.owl#"
  xml:base="http://www.hozo.jp/owl/TEdPM.owl#"
>
<owl:Ontology rdf:about="">
  <rdfs:comment>
    HOZO:OWL Export
  </rdfs:comment>
</owl:Ontology>

<owl:Class rdf:ID="RelationalConcept">
  <rdfs:label>RelationalConcept</rdfs:label>
</owl:Class>

<owl:ObjectProperty rdf:ID="hasPart">
  <rdfs:label>hasPart</rdfs:label>
</owl:ObjectProperty>

<owl:ObjectProperty rdf:ID="hasAttribute">
  <rdfs:label>hasAttribute</rdfs:label>
</owl:ObjectProperty>
```

ภาพที่ 4: ตัวอย่าง OWL Ontology (RDF/XML Source Code)

จากรูปที่ 3 แสดงตัวอย่าง OWL Ontology (RDF/XML Source Code) ของ TCSEdPM ออนไลน์ที่สร้างขึ้น

หลังจากนั้นใช้ระบบจัดการโปรแกรมประยุกต์ฐานความรู้ออนไลน์ (Ontology Application Management (OAM Framework User Tutorial) ของหน่วยปฏิบัติการวิจัยเทคโนโลยี

ภาษาธรรมชาติและความหมาย [3] สร้างการเชื่อมต่อกับระบบฐานข้อมูล MySQL โดยใช้โปรแกรม Navicat for MySQL เพื่อนำเข้าข้อมูลที่จัดเตรียมไว้ขึ้นสู่ระบบฐานข้อมูล และใช้โปรแกรม Configuration เพื่อเชื่อมต่อระบบฐานข้อมูลกับไฟล์ฐานความรู้ออนโทโลยี (DB-Ontology Mapping) ด้วยการกำหนดความสัมพันธ์ระหว่างฐานข้อมูลและฐานความรู้ออนโทโลยีเพื่อแปลงข้อมูล (Data Mapping)

กระบวนการสุดท้าย ซึ่งจะดำเนินการในขั้นตอนต่อไปของงานวิจัย คือ การนำ TCSEdPM Ontology ที่สร้างเสร็จแล้วไปประยุกต์ใช้เพื่อการพัฒนาในระบบในรูปแบบระบบสืบค้นข้อมูลเชิงความหมาย (Semantic Ontology Search) และระบบให้คำแนะนำ (Advisory System) ให้เป็นระบบที่มีความสามารถในการแนะนำข้อมูลสำหรับงานการบริหารบุคคลข้าราชการครูและบุคลากรทางการศึกษา

4. ออนโทโลยีที่ได้

TCSEdPM เป็นโดเมนออนโทโลยีที่มีเนื้อหาเฉพาะเรื่องการบริหารงานบุคคลข้าราชการครูและบุคลากรทางการศึกษา ที่มีลักษณะเป็นโครงสร้างลำดับชั้นแสดงถึงความสัมพันธ์ความสัมพันธ์แบบทั่วไปโดยใช้แลเบล (Label) และชื่อลำดับชั้นด้วยความสัมพันธ์แบบจัดเป็น (is-a) ประกอบด้วยแนวคิด (Concept) จำนวน 11 แนวคิด ดังนี้คือ 1) กรณีศึกษา 2) วิธีการบรรจุและแต่งตั้ง 3) ประเภทบุคลากร 4) ตำแหน่ง 5) ระดับวิทยฐานะ 6) สังกัด 7) คุณสมบัติการศึกษา 8) วิชาเอก 9) กระทรวง 10) บัญชีผู้สอบแข่งขันได้ 11) เอกสารอ้างอิง ซึ่งมีการกำหนดให้มีความสัมพันธ์กัน เช่น การพิจารณาบุคคลเพื่อบรรจุและแต่งตั้งในแต่ละตำแหน่งจะมีความสัมพันธ์กับคุณสมบัติประจำตำแหน่งและคุณสมบัติทั่วไปของผู้สมัคร เป็นต้น

5. สรุปผลการศึกษาวิจัย

งานวิจัยฉบับนี้ผู้วิจัยได้นำเสนอการออกแบบและพัฒนา TCSEdPM ออนไลน์งานการบริหารงานบุคคลข้าราชการครูและบุคลากรทางการศึกษา การนำโปรแกรม Hozo-Ontology Editor Version 5.2.36 ซึ่งเป็นโปรแกรมที่ผู้วิจัยไม่เคยมีประสบการณ์ในการใช้งานมาก่อน ทำให้ยากและเสียเวลามากในการศึกษาเพื่อที่จะพัฒนา TCSEdPM ออนไลน์ และการออกแบบออนโทโลยีเกี่ยวกับการบริหารงานบุคคล

ข้าราชการครูและบุคลากรทางการศึกษามีรายละเอียดมากมาย ทำให้ต้องปรับปรุงการออกแบบรายละเอียดของคลาสและความสัมพันธ์ระหว่างคลาสดังๆ ตลอดเวลา สำหรับ TCSEdPM ออนโทโลยีที่ได้ ขณะนี้ผู้วิจัยยังไม่ได้ดำเนินการประเมินประสิทธิภาพของออนโทโลยี ดังนั้นขั้นต่อไปของงานวิจัยจะเป็นการประเมินประสิทธิภาพการใช้งาน [13] ด้วยการพิจารณาค่าความถูกต้อง ค่าความแม่นยำ (Precision) และค่าความระลึก (Recall) และการวัดประสิทธิภาพโดยรวม (F-measure) ของผลการสืบค้นโดยผู้ใช้งาน และออนโทโลยีนี้จะถูกนำไปใช้ร่วมกับเทคนิคการให้เหตุผลด้วยกรณีศึกษา (CBR) เพื่อเป็นส่วนหนึ่งของการสร้างซอฟต์แวร์ (Software Construction) สำหรับระบบให้คำแนะนำการบริหารงานบุคคลข้าราชการครูและบุคลากรทางการศึกษา ทำให้ส่วนประมวลผลเชิงอนุมาน (Inference Engine) แนะนำข้อมูลให้กับผู้ใช้งานได้อย่างมีประสิทธิภาพมากขึ้นและถูกต้องมากที่สุด และเมื่อพัฒนาระบบให้คำแนะนำเสร็จแล้วจะทำการศึกษากายอมรับเทคโนโลยีใหม่เข้าไปใช้ในองค์กร (Technology Acceptance) โดยเก็บข้อมูลจากผู้ใช้งานระบบให้คำแนะนำในการปฏิบัติงานต่อไป

6. กิตติกรรมประกาศ

ขอขอบคุณ ดร.เทพชัย ทรัพย์นิธิ ดร.มารุต บุรณรัช และคณะนักวิจัย หน่วยปฏิบัติการวิจัยเทคโนโลยีภาษาธรรมชาติและภาษา ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ

ขอขอบคุณทุนอุดหนุนการวิจัยเพื่อทำวิทยานิพนธ์ของนักศึกษาในระดับบัณฑิตศึกษา บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ ประจำปีงบประมาณ 2556 ครั้งที่ 1 ที่สนับสนุนทุนการศึกษาให้แก่ ดวงพร เจริญอัมพร

เอกสารอ้างอิง

- [1] N.F. Noy and D.L. McGuinness, "Ontology Development 101: A Guide to Creating Your First Ontology," Stanford Knowledge System Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880, Stanford University, 2001.
- [2] V.C. Storey. et al., "An Ontology-Based Expert System for Database Design," Data and Knowledge Engineering, 28(1): pp. 31-46, 1998.
- [3] ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ, "แนะนำเครื่องมือการพัฒนาฐานความรู้เพื่อนำไปสู่การประยุกต์ใช้ข้อมูลเชิงความหมาย ห้องปฏิบัติการวิจัยเทคโนโลยีภาษาธรรมชาติและภาษา" เอกสารประกอบการอบรมเชิงปฏิบัติการ Semantic-based Knowledge Management Tools รุ่นที่ 13 วันที่ 25-27 มกราคม 2556 ณ อุทยานวิทยาศาสตร์ประเทศไทย, 2556.
- [4] สำนักงาน ก.ค.ศ., "Q&A เกี่ยวกับการบริหารงานบุคคลของ อ.ก.ค.ศ. เขตพื้นที่การศึกษา," กรุงเทพฯ: สำนักงานปลัดกระทรวงศึกษาธิการ กระทรวงศึกษาธิการ, 2555.
- [5] M.-H. Wang, et al., "Ontology-based multi agents for intelligent healthcare applications," Journal of Ambient Intelligence and Humanized Computing 1(2): 111-131, 2010.
- [6] LW. Lacy, "OWL: Representing: Information Using the Web Ontology Language," Canada: Trafford Publishing (uk) Ltd., 2005.
- [7] นกัศ สุขสม, มารุต บุรณรัช และ เทพชัย ทรัพย์นิธิ, "การพัฒนาออนโทโลยีสำหรับระบบให้คำแนะนำการบริโภคอาหารตามโภชนาการเฉพาะบุคคล," กรุงเทพฯ: ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ, 2551.
- [8] S. Izumi. Et al., "An Ontology-Based Advice System for Health And Exercise," Proceedings of the 10th IASTED International Conference on Internet and Multimedia Systems and Application, 2006.
- [9] AS Aparicio, OLM Farias, N dos Santos, "Applying Ontologies in the Integration of Heterogeneous Relational Databases," Australasian Ontology Workshop (AOW 2005), Sydney. Conferences in Research and Practice in Information Technology (CRPIT), 2005.
- [10] F. Hakimpour, A. Geppert, "Resolving Semantic Heterogeneity in Schema Integration: An Ontology Based Approach," Formal Ontology in Information Systems Archive Proceedings of the International Conference on Formal Ontology in Information Systems, Ogunquit, Maine, USA., 2001.
- [11] X. Jiang, AH. Tan, "Learning and Inferencing in user ontology for personalized Semantic Web search," Information sciences 179(16): 2794-2808, 2009.
- [12] มหาวิทยาลัยโอซากา (ออนไลน์), "โปรแกรม Hozo-Ontology Editor," สืบค้นจาก: <http://www.hozo.jp/>, 15 มกราคม 2557.
- [13] พยุง มีสัจ, วาทีน นุ้ยเพียร และมุสดี บุญรอด, "ระบบการค้นคืนเชิงความหมายจากงานวิจัยโดเมน Information System," การประชุมวิชาการระดับประเทศทางด้านเทคโนโลยีสารสนเทศ (NCIT2012), ระหว่างวันที่ 26-27 เมษายน 2555, ชะอำ:เพชรบุรี, หน้า 320-324, 2555.

**การออกแบบออนโทโลยีผลงานวิจัยครูและบุคลากรทางการศึกษา:
กรณีศึกษานักงานเขตพื้นที่การศึกษาประถมศึกษาอุบลราชธานี เขต 5**
**Design ontologies Technology Research of Teachers and Educational Personnel:
Case study of The Ubon Ratchathani Primary Educational Service Area Office5**

ฐิตาภัทร์ อุทการ (Titapat Utthakan)¹ สมปอง เวฬุวนาธร (Sompong Valuvanathorn)²

และชัชวิน นามมัน (Chatchawin Namman)³

ภาควิชาคณิตศาสตร์ สถิติ และคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยอุบลราชธานี

¹yipzaub@gmail.com, ²scsompva@ubu.ac.th, ³scchatma@ubu.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อออกแบบออนโทโลยีเกี่ยวกับผลงานวิจัยของครูและบุคลากรทางการศึกษากรณีศึกษานักงานเขตพื้นที่การศึกษาประถมศึกษาอุบลราชธานี เขต 5 ซึ่งได้สร้างเป็นฐานความรู้ออนโทโลยี จากออนโทโลยีที่ได้สามารถใช้ข้อมูลร่วมกันและเชื่อมโยงความสัมพันธ์กันได้อย่างอัตโนมัติ จึงเป็นประโยชน์เพื่อการสืบค้นได้ข้อมูลที่ตรงกับความต้องการและความหมายอย่างแท้จริง และสามารถนำไปประยุกต์ใช้ในการรวบรวมองค์ความรู้อื่นๆ ให้เกิดประโยชน์และมีประสิทธิภาพต่อไป

คำสำคัญ: ออนโทโลยี ภาษาเชิงความหมาย เว็บเชิงความหมาย

Abstract

This research aims to design ontology about teachers and educational officers researches; a case study of The Ubon Ratchathani Primary Educational Service Area Office 5. The ontology was designed to be knowledge based ontology which could autonomously share and link information therefore it available for gathering and reviewing meaningful data. Moreover, the tool could be applied to other kind of knowledge searching for the more proficiency and accuracy outputs.

Keyword: Ontology, OWL, Semantic Web

1. บทนำ

ปัจจุบันการเข้าถึงข้อมูลความรู้ การแลกเปลี่ยนข้อมูลข่าวสารสารสนเทศนั้น มีความสะดวกและรวดเร็วมากยิ่งขึ้น

โดยอินเทอร์เน็ตนั้นมีบทบาทสำคัญในการสืบค้นข้อมูลความรู้ข่าวสารสารสนเทศจากแหล่งข้อมูลที่มีอยู่บนเครือข่ายอินเทอร์เน็ตมาใช้งานหรือใช้ประโยชน์ร่วมกันได้อย่างกว้างขวางในขณะเดียวกันแหล่งข้อมูลที่ว่านั้น นับวันได้เพิ่มจำนวนมากขึ้นเรื่อยๆ มากมายอย่างไม่มีการจำกัด จึงเป็นเหตุผลให้มีการกำหนดแนวคิดและทิศทางการพัฒนาไปถึงมาตรฐานสำหรับข้อมูลบนเว็บเชิงความหมาย ที่มีลักษณะการเชื่อมโยงข้อมูลในรูปแบบฐานความรู้หรือออนโทโลยี (Ontology) มีการแบ่งปันและใช้ข้อมูลร่วมกันได้อย่างมีประสิทธิภาพ นำไปสู่การพัฒนาโปรแกรมตัวแทน (Intelligent Agent) ที่มีความอัจฉริยะมากยิ่งขึ้น ซึ่งจะช่วยแก้ปัญหาความซ้ำซ้อนหลากหลายของข้อมูล การสืบค้นข้อมูลที่ได้อาจไม่ตรงกับความต้องการ เนื่องจากข้อมูลที่มีอยู่อย่างหลากหลาย งานวิจัยนี้จึงมีวัตถุประสงค์ในการออกแบบออนโทโลยี เพื่อนำไปประยุกต์ใช้ในการสืบค้นข้อมูลเกี่ยวกับผลงานวิจัยครูและบุคลากรทางการศึกษาของสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุบลราชธานี เขต 5 เป็นการจัดระเบียบให้กับข้อมูลที่มีความหมายให้เชื่อมโยงกันสามารถทำการสืบค้นข้อมูลได้ตรงกับที่ต้องการอย่างฉลาดและดียิ่งขึ้น อีกทั้งเพิ่มประสิทธิภาพและความสามารถในการเข้าถึงข้อมูลการค้นหาข้อมูลเป็นไปอย่างง่ายสะดวกและรวดเร็ว และที่สำคัญคือข้อมูลที่ได้อาจตรงกับความต้องการหรือความหมายอย่างแท้จริง

2. วรรณกรรมที่เกี่ยวข้อง

2.1 ออนโทโลยี

ออนโทโลยีคือการนิยามความหมายของแนวคิด(Concepts) หรือการอธิบายเชิงความหมายเพื่อความเข้าใจร่วมกันในขอบเขตความรู้ (Domain) เรื่องใดเรื่องหนึ่ง ให้สอดคล้องตรงกัน โดยใช้แนวคิดเดียว ในการลดหรือตัดทอนแนวคิดที่สับสน โดยมีการกำหนดคุณสมบัติ (Properties) และลักษณะความสัมพันธ์ (Relationships) เพื่อสร้างความหมายในลักษณะความสัมพันธ์ที่ถูกต้อง หรือระหว่างแนวคิดกับคุณสมบัติในขอบเขตความรู้ เพื่อสร้างองค์ความรู้ในด้านนั้นๆ โดยแสดงเป็นโครงสร้างความสัมพันธ์ที่ชัดเจน และสนับสนุนการแลกเปลี่ยน การสืบค้นข้อมูลสารสนเทศ ที่มีความสำคัญต่อการพัฒนาระบบฐานความรู้ (Knowledge Based Systems) เพราะออนโทโลยีนั้นสามารถแบ่งปัน ใช้ความรู้ร่วมกันและนำมาใช้ใหม่ได้ [1]

2.2 ภาษาเชิงความหมาย

ภาษาเชิงความหมาย (OWL) เป็นองค์ประกอบหนึ่งในเว็บเชิงความหมาย (Semantic Web) ที่ใช้ในการบรรยายข้อมูลเชิงความหมาย โดยการกำหนดโครงสร้างข้อมูลในลักษณะลำดับชั้น และอธิบายข้อมูลที่มีความสัมพันธ์ในระบบฐานข้อมูลได้ ทำให้ข้อมูลที่ถูกแทนที่นั้นมีความหมายมากยิ่งขึ้น ซึ่งจะอยู่ในรูปคลาส คุณสมบัติของคลาสและความสัมพันธ์ระหว่างคลาส เพื่ออธิบายเอนทิตีและความสัมพันธ์ต่างๆ [2]

2.3 เว็บเชิงความหมาย

เว็บเชิงความหมาย เป็นเว็บที่ได้ถูกออกแบบให้เป็นพื้นที่สำหรับแหล่งข้อมูล โดยมีเป้าหมายที่ว่าไม่ได้มีเพียงแค่ประโยชน์เพื่อสื่อสารกันระหว่างมนุษย์กับมนุษย์เท่านั้น แต่เครื่องคอมพิวเตอร์ต้องสามารถมีปฏิสัมพันธ์และให้ความช่วยเหลือได้ด้วย แต่ปัญหาคือข้อมูลส่วนใหญ่ที่ได้จากการค้นหาข้อมูลบนเว็บนั้น ยังต้องอาศัยการตัดสินใจของมนุษย์อยู่ แม้ว่าข้อมูลบนเว็บนั้นจะผ่านกระบวนการวิเคราะห์หรือนิยามไว้ดีเพียงใดก็ตาม เพื่อแก้ปัญหานี้ จึงได้มีการพัฒนาเว็บเชิงความหมายขึ้นมาใช้กับเว็บที่มีอยู่ เพื่อช่วยให้คอมพิวเตอร์มีความอัจฉริยะมากขึ้นในกระบวนการสืบค้นรวบรวมและการจัดการกับสารบัญเว็บและเข้าใจความหมายได้เหมือนมนุษย์ [3]

แนวคิดหลักเว็บเชิงความหมาย [4] มีดังนี้

1.การสร้างโมเดล การจัดเรียงลำดับ จัดแบ่งหรือจำแนก เพื่ออธิบายถึงความรู้ต่างๆ ที่เป็นสากลด้วยคำอธิบายย่อหรือคำสั้นๆ

2.การประมวลผลความรู้ โดยการใช้โมเดลความรู้ และเข้ารหัสข้อมูลความรู้เหล่านี้ด้วยเครื่องมือที่สามารถสร้างข้อสรุปที่มีความหมายหรือคอมพิวเตอร์สามารถรับรู้และเข้าใจได้

3.การแลกเปลี่ยนข้อมูลที่ซับซ้อน เพื่อที่จะบูรณาการความรู้ สร้างความเชื่อมโยงเป็นเครือข่าย บนฐานความรู้ที่มีขนาดใหญ่

การพัฒนาเว็บไปสู่เว็บเชิงความหมาย จึงจำเป็นที่จะต้องมีการออกแบบโมเดลหรือโครงสร้างเพื่อสร้างเป็นฐานข้อมูล นั่นก็คือฐานความรู้หรือออนโทโลยีที่เครื่องคอมพิวเตอร์สามารถเข้าใจได้อย่างมีความหมาย และมีกระบวนการวิธีการในการแลกเปลี่ยนข้อมูลเพื่อบูรณาการร่วมกันและเชื่อมโยงข้อมูลเข้าด้วยกัน โดยยึดหลักการและมาตรฐานตามที่กำหนด

2.2 งานวิจัยที่เกี่ยวข้อง

รวิกานต์ [5] ได้นำประโยชน์จากคุณลักษณะความสามารถของเว็บเชิงความหมายมาประยุกต์ใช้ในการสืบค้นความเชี่ยวชาญของนักวิจัยในประเทศไทยโดยการสร้างออนโทโลยีเพื่อจำแนกและวิเคราะห์ทักษะความเชี่ยวชาญของนักวิจัยและพัฒนาระบบสืบค้นผู้เชี่ยวชาญของนักวิจัย ซึ่งผลการประเมินระบบให้ความแม่นยำในการสืบค้นอยู่ในประสิทธิภาพดี

อนุวัฒน์และนัทรตระกูล [6] ได้นำเทคโนโลยี เว็บเชิงความหมาย (Semantic Web Service) มาประยุกต์ใช้ในการค้นหาในเว็บไซค์ทั่วไปโดยใช้ภาษาของออนโทโลยี (Ontology Language) ในการอธิบายรูปแบบโครงสร้างความสัมพันธ์ของข้อมูล และได้ให้รายละเอียดของค่าลงไปเพื่อช่วยในการค้นหา โดยการเชื่อมโยงความสัมพันธ์ของคำให้เกิดความหมายใกล้เคียงและถูกต้องทำให้ผู้ใช้งานที่มีความต้องการในการค้นหาที่สมบูรณ์แบบและตรงตามความต้องการมากยิ่งขึ้น

Sim และ Crowder [7] ได้พัฒนาระบบสืบค้นผู้เชี่ยวชาญในองค์กร โดยสร้างโมเดลความเชี่ยวชาญ (Expert Model) จากการใช้วิเคราะห์ความสัมพันธ์ของข้อมูลต่างๆ ในเอกสารที่เกี่ยวข้องกับบุคลากรในองค์กร เช่น เอกสารงานวิจัยต่างๆ ของบุคลากร และข้อมูลการทำงานในองค์กรแล้วพิจารณา

ความเชี่ยวชาญจากการจับคู่ คำสำคัญ (Keyword) ในหัวข้อ และส่วนเนื้อหาของเอกสารเหล่านั้นแล้วกำหนดค่าคะแนน แสดงความเชี่ยวชาญของแต่ละบุคคล

Zongmin และ Huaiqing [8] ได้เสนอวิธีการแยกแยะการจัดการความรู้และการจัดการข้อมูลโดยให้ความสำคัญอย่างยิ่งกับความหมาย เพราะมีการกำหนดวิธีการกึ่งอัตโนมัติของการสร้างข้อมูลดังกล่าวได้อย่างรวดเร็วด้วยประสิทธิภาพของออนโทโลยี และได้นำเสนอวิธีการใหม่ในการแยกข้อมูลอัตโนมัติ ที่มีประสิทธิภาพมากกว่าซอฟต์แวร์ธรรมดา การสร้างโดเมนออนโทโลยี หรือฐานความรู้ ซึ่งไม่ได้มีเพียงฐานวิศวกรรมความรู้ (Engineering Ontology) แต่ยังสามารถประยุกต์ใช้ในสาขาอื่นได้อีกมาก โดยใช้แผนภาพในการออกแบบ เพื่อแยกแยะและเป็นตัวแทนแนวความคิดของโดเมนการจัดการความรู้และการจัดการข้อมูลอย่างชัดเจนร่วมกัน

จากการศึกษาผลการวิจัย ซึ่งแสดงให้เห็นว่าการสร้างออนโทโลยีหรือฐานความรู้มาประยุกต์ใช้งานนั้น สามารถลดขั้นตอนหรือช่วงเวลาในการค้นหาข้อมูลที่ต้องการตามวัตถุประสงค์ของแต่ละบุคคลได้อย่างมีประสิทธิภาพและแสดงผลได้ชัดเจนถึงความแตกต่างและประโยชน์ของการนำข้อมูลเหล่านั้นมาใช้งานได้อย่างสะดวกง่ายดายมากยิ่งขึ้น

3. วิธีการดำเนินการวิจัย

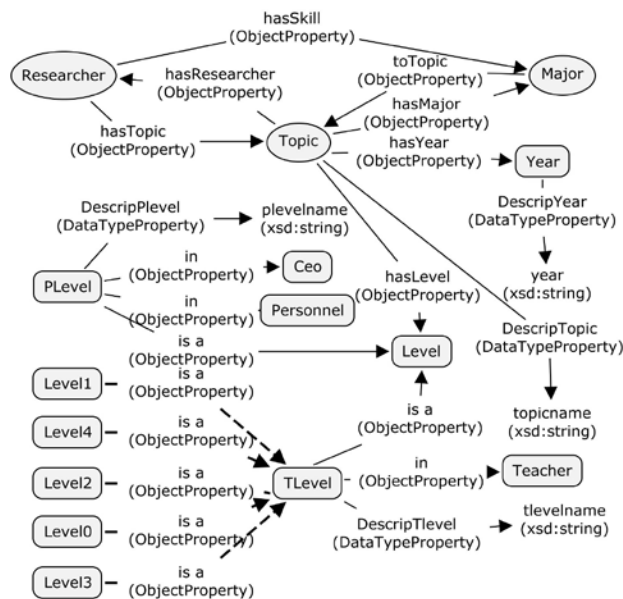
ผู้วิจัยได้ดำเนินการตามขั้นตอน ดังต่อไปนี้

3.1 รวบรวมข้อมูล

เก็บรวบรวมข้อมูลผลงานทางวิชาการ งานวิจัยของข้าราชการครูและบุคลากรทางการศึกษาของสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุบลราชธานี เขต 5 ได้แก่ ผลงานประเมินวิทยฐานะ ผลงานประเมินเลื่อนระดับตำแหน่งที่สูงขึ้น และงานวิจัย ซึ่งเป็นแหล่งข้อมูลความรู้ภายในองค์กร

3.2 ออกแบบโครงสร้างออนโทโลยี

จากการเก็บรวบรวมข้อมูลผลงานทางวิชาการ งานวิจัยของข้าราชการครูและบุคลากรทางการศึกษาดังกล่าว จึงทำการออกแบบโครงสร้างออนโทโลยี ดังภาพที่ 1 ซึ่งสามารถอธิบายความสัมพันธ์ได้ในหัวข้อที่ 3.3



ภาพที่ 1: ตัวอย่างโครงสร้างออนโทโลยีผลงานวิจัย

3.3 วิเคราะห์ข้อมูลออนโทโลยี

หลังจากที่ได้ออกแบบโครงสร้างออนโทโลยีแล้วได้กำหนดออนโทโลยีหลัก ดังนี้

- 1) ออนโทโลยีผลงานวิจัย ได้มีการวิเคราะห์และจัดเก็บเป็นรูปแบบของคลาส ดังนี้

ตารางที่ 1: รายละเอียดของ Class ในออนโทโลยีผลงานวิจัย

ลำดับ	Class	รายละเอียด
1	Topic	คลาสแทนข้อมูลหัวข้อผลงานวิจัย
2	Level	คลาสแทนข้อมูลระดับผลงานวิจัย
3	TLevel	คลาสแทนข้อมูลผลงานวิจัยของข้าราชการครู
4	PLevel	คลาสแทนข้อมูลผลงานวิจัยของบุคลากรทางการศึกษา
5	Level0	คลาสแทนข้อมูลผลงานวิจัยของข้าราชการครูระดับปฐมวัย
6	Level1	คลาสแทนข้อมูลผลงานวิจัยของข้าราชการครูระดับช่วงชั้นที่ 1
7	Level2	คลาสแทนข้อมูลผลงานวิจัยของข้าราชการครูระดับช่วงชั้นที่ 2
8	Level3	คลาสแทนข้อมูลผลงานวิจัยของข้าราชการครูระดับช่วงชั้นที่ 3
9	Level4	คลาสแทนข้อมูลผลงานวิจัยของข้าราชการครูระดับช่วงชั้นที่ 4
10	Year	คลาสแทนข้อมูลปีที่จัดทำ

2) ออนโทโลยีผู้วิจัย ได้มีการวิเคราะห์และจัดเก็บเป็นรูปแบบของคลาส ดังนี้

ตารางที่ 2: รายละเอียดของ Class ในออนโทโลยีผู้วิจัย

ลำดับ	Class	รายละเอียด
1	Researcher	คลาสแทนข้อมูลผู้วิจัย
2	Position	คลาสแทนข้อมูลตำแหน่งของผู้วิจัย
3	Degree	คลาสแทนข้อมูลระดับการศึกษา
4	Contact	คลาสแทนข้อมูลที่อยู่ติดต่อ
5	Ceo	คลาสแทนข้อมูลตำแหน่งผู้บริหาร
6	Personnel	คลาสแทนข้อมูลตำแหน่งบุคลากรทางการศึกษา
7	Teacher	คลาสแทนข้อมูลตำแหน่งครู
8	Bachelor	คลาสแทนข้อมูลปริญญาตรี
9	Master	คลาสแทนข้อมูลปริญญาโท
10	Doctor	คลาสแทนข้อมูลปริญญาเอก
11	SubDistrict	คลาสแทนข้อมูลตำบล
12	District	คลาสแทนข้อมูลอำเภอ
13	Province	คลาสแทนข้อมูลจังหวัด
14	Postcode	คลาสแทนข้อมูลรหัสไปรษณีย์
15	Tel	คลาสแทนข้อมูลหมายเลขโทรศัพท์

3) ออนโทโลยีหมวดหมู่ ได้มีการวิเคราะห์และจัดเก็บเป็นรูปแบบของคลาส ดังนี้

ตารางที่ 3: รายละเอียดของ Class ในออนโทโลยีหมวดหมู่ผลงานวิจัย

ลำดับ	Class	รายละเอียด
1	Major	คลาสแทนข้อมูลหมวดหมู่ผลงานวิจัย
2	Thai	คลาสแทนข้อมูลหมวดหมู่ภาษาไทย
3	Math	คลาสแทนข้อมูลหมวดหมู่คณิตศาสตร์
4	Science	คลาสแทนข้อมูลหมวดหมู่วิทยาศาสตร์
5	Social	คลาสแทนข้อมูลหมวดหมู่สังคมศึกษาศาสตร์และวัฒนธรรม
6	Health	คลาสแทนข้อมูลหมวดหมู่สุขภาพและพลศึกษา
7	Art	คลาสแทนข้อมูลหมวดหมู่ศิลปะ
8	Techno	คลาสแทนข้อมูลหมวดหมู่การงานอาชีพและเทคโนโลยี
9	Foreign	คลาสแทนข้อมูลหมวดหมู่ภาษาต่างประเทศ
10	Kid	คลาสแทนข้อมูลหมวดหมู่ปฐมวัย
11	Etc	คลาสแทนข้อมูลหมวดหมู่อื่นๆ

จากออนโทโลยีหลักที่ได้จึงกำหนดความสัมพันธ์ระหว่างคลาสโดยรายละเอียดของ Object Property ในแต่ละออนโทโลยีเป็นการบอกคุณสมบัติการเข้าถึงข้อมูลระหว่างคลาสได้ โดยผ่านคุณสมบัติความสัมพันธ์เชิงความหมาย

ตารางที่ 4: รายละเอียดของ Object Property ผลงานวิจัย

ลำดับ	Object Property	รายละเอียด
1	hasMajor	แทนความสัมพันธ์ระหว่าง คลาส Topic และ คลาส Major
2	hasResearcher	แทนความสัมพันธ์ระหว่าง คลาส Topic และ คลาส Researcher
3	hasLevel	แทนความสัมพันธ์ระหว่าง คลาส Topic และ คลาส Level
4	hasYear	แทนความสัมพันธ์ระหว่าง คลาส Topic และ คลาส Year
5	is a	แทนความสัมพันธ์ระหว่าง คลาส Level กับ คลาส PLevel และคลาส TLevel กับคลาส Level0 คลาส Level1 คลาส Level2 คลาส Level3 คลาส Level4
6	in	แทนความสัมพันธ์ระหว่าง คลาส TLevel กับคลาส Teacher คลาส PLevel กับคลาส Ceo และคลาส Personnel
7	hasTopic	แทนความสัมพันธ์ระหว่าง คลาส Researcher และคลาส Topic
8	toTopic	แทนความสัมพันธ์ระหว่าง คลาส Major และ คลาส Topic

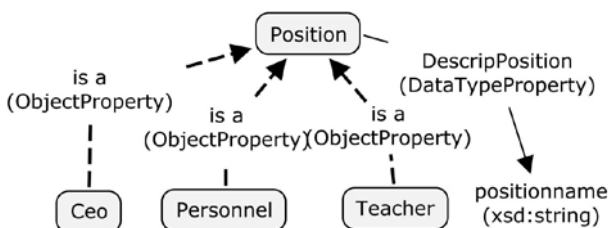
ตารางที่ 5: รายละเอียดของ DataType Property ผลงานวิจัย

ลำดับ	DataType Property	ชนิดข้อมูล	รายละเอียด
1	DescripTopic	String	ชื่อผลงาน
2	DescripTlevel	String	ชื่อผลงานของข้าราชการครู
3	DescripPlevel	String	ชื่อผลงานของบุคลากรทางการศึกษา
4	DescripYear	String	ปีที่จัดทำ

ส่วน DataType Property เป็นการกำหนดคุณสมบัติของข้อมูลที่อยู่ในคลาสนั้นๆ เช่น คลาส Topic จะมี DataType Property คือ DescripTopic DescripTlevel DescripPlevel และ DescripYear

3.4 จัดกลุ่มข้อมูลด้วยภาษาเชิงความหมาย

เมื่อกำหนดรูปแบบของคลาส (Class) ความสัมพันธ์ (Object Property) และ คุณสมบัติของคลาส (Data Type Property) จากนั้นได้จัดกลุ่มของข้อมูลโดยการกำหนดคลาสย่อยที่มีรายละเอียดตามคลาสหลัก ซึ่งต้องมีความสัมพันธ์กัน เช่น คลาส Position มีคลาสย่อย คือ คลาส Ceo คลาส Personnel คลาส Teacher ดังภาพ



ภาพที่ 2: ตัวอย่างโครงสร้างออนโทโลยีตำแหน่งของผู้วิจัย

โดยมีการกำหนดรูปแบบภาษาเชิงความหมาย ดังนี้

```
<Topicmain:Ceo rdf:ID="Ceo1">
  <Topicmain:subClassOf rdf:resource="#Position"/>
</Topicmain:Ceo>
<Topicmain:Personnel rdf:ID="Personnel2">
  <Topicmain:subClassOf rdf:resource="#Position"/>
</Topicmain:Personnel>
<Topicmain:Teacher rdf:ID="Teacher3">
  <Topicmain:subClassOf rdf:resource="#Position"/>
</Topicmain:Teacher>
```

ภาพที่ 3: ตัวอย่างการกำหนดคลาสย่อย

4. ผลการดำเนินงาน

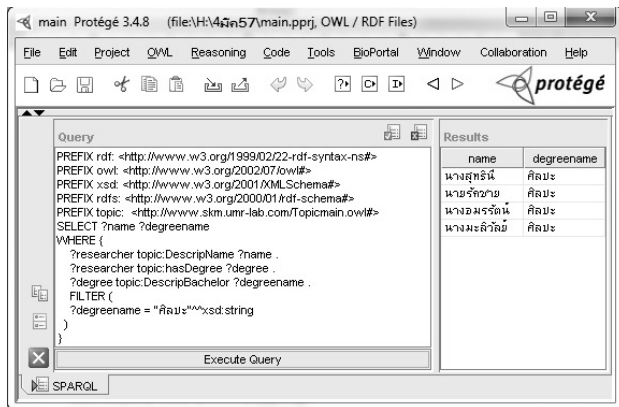
เมื่อทำการจัดกลุ่มของข้อมูลเรียบร้อยแล้ว จึงได้ออนโทโลยีในรูปแบบภาษาเชิงความหมาย ดังนี้

```
<?xml version="1.0"?>
<rdf:RDF
  xmlns="http://localhost/ontology/Researcher.owl#"
  xml:base="http://localhost/ontology/Researcher.owl"
  xmlns:owl2xml="http://www.w3.org/2006/12/owl2-xml#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:Topicmain="http://localhost/ontology/Topicmain.owl#"
  xmlns:Topic="http://localhost/ontology/Topic.owl#"
  xmlns:Major="http://localhost/ontology/Major.owl#"
  xmlns:Level="http://localhost/ontology/Level.owl#"
  xmlns:Position="http://localhost/ontology/Position.owl#"
  xmlns:Degree="http://localhost/ontology/Degree.owl#"
  xmlns:Contact="http://localhost/ontology/Contact.owl#"
  xmlns:Skill="http://localhost/ontology/Skill.owl#"
  xmlns:daml="http://www.daml.org/2001/03/daml+oil#">
  <Topicmain:Researcher rdf:about="#R01">
    <Topicmain:Name rdf:datatype="xsd:string">
      นางอังชรา</Topicmain:Name>
    <Topicmain:LastName rdf:datatype="xsd:string">
      จันทร์ระน้อย</Topicmain:LastName>
    <Topicmain:Email
      rdf:datatype="xsd:string">Email@hotmail.com</Topic
      main:Email>
    <Topicmain:hasTopic rdf:resource="&Topic;T01"/>
    <Topicmain:hasSkill rdf:resource="&Skill;Art"/>
    <Topicmain:hasContact rdf:resource="&Contact;C01"/>
    <Topicmain:hasDegree rdf:resource="&Degree;Be01"/>
    <Topicmain:hasPosition
      rdf:resource="&Position;Teacher3"/>
    </Topicmain:Researcher>
    <Topicmain:Researcher rdf:about="#R02">
      <Topicmain:Name rdf:datatype="xsd:string">นางสุทธิณี
    </Topicmain:Name>
    <Topicmain:LastName rdf:datatype="xsd:string">นิชรรัตน์
    </Topicmain:LastName>
    <Topicmain:Email
      rdf:datatype="xsd:string">Email2@hotmail.com
    </Topicmain:Email>
    <Topicmain:hasTopic rdf:resource="&Topic;T02"/>
    <Topicmain:hasSkill rdf:resource="&Skill;Art"/>
    <Topicmain:hasContact rdf:resource="&Contact;C02"/>
    <Topicmain:hasDegree rdf:resource="&Degree;Be02"/>
    <Topicmain:hasPosition
      rdf:resource="&Position;Teacher3"/>
    </Topicmain:Researcher>
  </rdf:RDF>
```

ภาพที่ 4: ตัวอย่างออนโทโลยีของผู้วิจัย

จากภาพที่ 4 คลาส Researcher มีข้อมูลที่เป็น Data Type Property ที่แสดงถึงตัวอย่างข้อมูล Instances หรือ Individual คือ Name LastName Email ส่วน Property ที่อธิบายความสัมพันธ์ของข้อมูล ประกอบด้วย hasTopic ซึ่งเชื่อมโยงข้อมูลกับ คลาส Topic ที่มี Data Type Property เป็น DescripTopic ชนิดข้อมูลเป็น String คือ ชื่อผลงาน ตามตารางที่ 5 ซึ่งระบุรายละเอียดของ Data Type Property แต่ละตัวที่มีความสัมพันธ์เชื่อมโยงกันนั่นเอง

ในการทดสอบออนโทโลยีที่สร้างขึ้น ผู้วิจัยได้ใช้ภาษา SPARQL ในการค้นหาข้อมูลบนโปรแกรม Protege ดังแสดงได้ในภาพที่ 5



ภาพที่ 5: ตัวอย่างการค้นหาข้อมูลด้วยภาษา SPARQL

จากภาพที่ 5 เป็นตัวอย่างการค้นหาข้อมูล ด้วยภาษา SPARQL โดยการกำหนดเงื่อนไขในการค้นหาข้อมูล เช่น ต้องการค้นหาข้อมูลผู้วิจัยที่มีระดับการศึกษาปริญญาตรีในสาขาศิลปะ จึงกำหนดตัวแปรที่จะแสดงผล คือ name (ชื่อผู้วิจัย) และ degreeName (สาขาศิลปะ) โดยมี Object Property ที่ใช้อธิบายความสัมพันธ์ของข้อมูล คือ hasDegree ซึ่งเชื่อมโยงข้อมูลกันระหว่าง คลาส Researcher กับ คลาส Degree จึงได้ผลลัพธ์ที่เป็นรายชื่อของผู้วิจัยที่จบปริญญาสาขา ศิลปะ ดังภาพ

5. สรุป

งานวิจัยนี้ได้สร้างออนโทโลยี โดยการออกแบบโมเดลเพื่อแทนความรู้เกี่ยวกับผลงานวิจัยของครูและบุคลากรทางการศึกษา ซึ่งได้เป็นฐานความรู้ภายในองค์กร เพื่อนำไปประยุกต์ใช้พัฒนาบนเว็บเชิงความหมาย และยังสามารถนำรูปแบบขั้นตอนและวิธีการในส่วนของการสร้างและออกแบบออนโทโลยีไปพัฒนาหรือต่อยอดเป็นฐานความรู้ด้านอื่นๆ เพื่อนำไปใช้ประโยชน์ในการสืบค้นข้อมูล และได้ข้อมูลที่มีประสิทธิภาพอย่างมีความหมาย

การออกแบบออนโทโลยี จึงเป็นสิ่งสำคัญในการสร้างออนโทโลยีสำหรับขอบเขตความรู้เรื่องใดเรื่องหนึ่ง ที่ต้องการกำหนดความสัมพันธ์ของแนวคิดและคุณสมบัติ เพื่อให้ได้

ความหมายตรงกับความต้องการอย่างถูกต้อง ตามโครงสร้างและการออกแบบออนโทโลยี และอยู่บนพื้นฐานแนวคิดที่เข้าใจตรงกัน สามารถนำไปประยุกต์ใช้กับองค์ความรู้ด้านอื่นๆ เพื่อใช้งานอย่างหลากหลายต่อไปในอนาคต

เอกสารอ้างอิง

- [1] Noy, Natalya and McGuinness, Deborah L, 2001, "Ontology Development 101: A Guide to Creating Your First Ontology". Available From: www.ksl.stanford.edu/people/dlm/papers/ontology-tutorial-noy-mcguinness.pdf
- [2] Smith MK, Welty C, McGuinness DL., 2004, "OWL Web Ontology Language Guide". Available From: <http://www.w3.org/TR/owlguide/index.html>
- [3] Tim Berners-Lee,1998, "A roadmap to the semantic web" Available From : <http://www.w3.org/DesignIssues/Semantic.html>
- [4] Pascal Hitzler, "Foundations of Semantic Web Technologies" CRC Press, 2001.
- [5] รวิภรณ์ ปันณะรัต, "การประยุกต์เว็บเชิงความหมายในการสืบค้นความเชี่ยวชาญของนักวิจัย"สารนิพนธ์ปริญญาโท สาขาวิชาเทคโนโลยีสารสนเทศ บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ,2552.
- [6] อนุวัฒน์ โยวนุตร และ นัทรชระกูด สมบัติธีระ, "การค้นหาข้อมูลด้วยการประยุกต์ใช้เว็บเชิงความหมาย," *The 3rd National CS&ICT Conference (NCSICT 2009)*, Oct 193-212, 2009.
- [7] Y. Sim and R. Crowder, "Evaluation of an Approach to Expertise Finding," *5th International Conference on Practical Aspects of Knowledge Management, Vienna, Austria, vol. 3336, pp.141-152.*2004.
- [8] Zongmin & Huaqing. "The Semantic Web for Knowledge and Data Management: Technologies and Practices Ontology" *A Review of Fuzzy Models for the Semantic Web, Hershey: The United States of America.*2009

การให้คำแนะนำทรัพยากรการเรียนรู้ส่วนบุคคลบนพื้นฐานของออนโทโลยี สำหรับผู้เรียนบกพร่องทางการเรียนรู้

An Ontology-based Personalized Learning Resource Recommendation for Learning Disability Students

วรชาติ สุขกระ (Worachat Sukra)¹ และ นพมาศ ปักเข็ม (Noppamas Pukkhem)²
สาขาวิชาคอมพิวเตอร์และเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยทักษิณ
¹Maxy2cutter@gmail.com, ²noppamas@tsu.ac.th.com

บทคัดย่อ

บทความนี้นำเสนอการออกแบบ พัฒนาและทดสอบการให้คำแนะนำทรัพยากรการเรียนรู้ส่วนบุคคลบนฐานความรู้ออนโทโลยีสำหรับผู้เรียนบกพร่องทางการเรียนรู้ เพื่อช่วยให้ผู้บกพร่องทางการเรียนรู้ ครูผู้สอน หรือผู้ที่เกี่ยวข้องสามารถเลือกใช้สื่อการเรียนรู้ที่เหมาะสมกับผู้บกพร่องทางการเรียนรู้ส่วนบุคคลได้ ในงานวิจัยนี้ได้พัฒนาออนโทโลยีที่แสดงถึงความสัมพันธ์กันระหว่างผู้เรียนบกพร่องทางการเรียนรู้และสื่อการเรียนรู้ที่เหมาะสม (ontoLRforLD) ผ่านการออกแบบตามหลักวงจรการออกแบบออนโทโลยี (Ontology Life Cycle) และแสดงผลในรูปแบบของเอกสาร RDF ด้วยโปรแกรม Protégé โดยใช้ภาษา Sparql Query ในการสอบถาม จากนั้นนำฐานความรู้ดังกล่าวมาเชื่อมต่อกับโปรแกรมประยุกต์ที่พัฒนาด้วยภาษา PHP เพื่อให้ระบบสามารถประมวลผลการแนะนำทรัพยากรการเรียนรู้การเรียนรู้ได้อย่างอัตโนมัติ โดยจากการประเมินความเหมาะสมฐานความรู้ออนโทโลยีทรัพยากรการเรียนรู้สำหรับผู้เรียนบกพร่องทางการเรียนรู้จากผู้เชี่ยวชาญได้ค่าเฉลี่ย (\bar{x}) = 4.16 ส่วนเบี่ยงเบน (S.D) = 0.11 ซึ่งถือว่าอยู่ในเกณฑ์ดี

คำสำคัญ: การให้คำแนะนำ ทรัพยากรการเรียนรู้ ความบกพร่องทางการเรียนรู้ ออนโทโลยี

Abstract

This paper proposes a process for designing, developing and testing of an ontology-based personalized learning resource recommendation for learning disabled

student. With support, learning disability students, teachers or a person who related to the students can select the most suitable individual learning resources. In this work, we develop an ontology called "ontoLRforLD", which represents the relationship between a disability student and a learning resource according to the ontology development life cycle. The ontology is implemented and converted into RDF format using Protégé ontology editor. For query processing, the Sparql query language will be used. The PHP user interface is developed for connecting the ontology and web-based application, so the application can be processed automatically. According to the ontoLRforLD knowledge-based evaluation by experts, the mean is 4.16 and S.D. is 0.11. That results show that ontology is a good level.

Keywords: Recommendation, Learning Resource, Learning Disability, Ontology

1. บทนำ

ในปัจจุบันสื่อการเรียนรู้ในรูปแบบดิจิทัลและแบบทั่วไปมีปริมาณเพิ่มมากขึ้นส่งผลให้ผู้เรียนประสบปัญหาในการเลือกสื่อการเรียนรู้ดังกล่าวที่ตรงกับความต้องการหรือความต้องการที่แท้จริง ทั้งนี้ต้องสอดคล้องกับการศึกษาซึ่งเป็นสิทธิของมนุษยชนจะต้องได้รับ เด็กที่มีความผิดปกติจำเป็นต้องมีเทคโนโลยีการสื่อสารเพื่อสามารถเข้าถึงข้อมูลได้อย่างตรงจุด [1] โดยเฉพาะผู้เรียนที่มีความบกพร่องทางการเรียนรู้หรือกลุ่มเด็กแอลดี (LD: Learning Disability child)

ซึ่งอาจปะปนอยู่ในกลุ่มผู้เรียนปกติทั่วไป ในบางครั้งผู้เรียนหรือครูผู้สอนอาจไม่ทราบว่าตัวนักเรียนมีลักษณะเข้าข่ายผู้เรียนบกพร่องทางการเรียนรู้ เพราะโดยทั่วไปเด็กจะมีลักษณะภายนอกที่เป็นปกติ และสามารถเรียนในบางรายวิชาได้อย่างปกติ แต่จะมีปัญหาเฉพาะบางรายวิชาที่จำเป็นจะต้องเพิ่มทักษะกระบวนการเรียนรู้ให้กับเด็กบกพร่องทางการเรียนรู้

ระบบที่ออกแบบมาแล้วไม่ตอบสนองความต้องการของผู้เรียนบกพร่องทางการเรียนรู้แสดงว่าการออกแบบระบบนั้นไม่ได้คาดหวังและตระหนักถึงความบกพร่องของเด็กผู้เรียนบกพร่องอย่างตรงจุด ซึ่งส่งผลต่อความรู้ที่ได้รับไม่เต็มที่ ผู้วิจัยจึงได้ทำการศึกษาถึงแนวคิดที่จะพัฒนาระบบให้คำแนะนำ (Recommendation System) ทรัพยากรการเรียนรู้สำหรับผู้เรียนบกพร่องทางการเรียนรู้ โดยรวบรวมข้อมูลเป็นระบบฐานความรู้ประยุกต์ใช้กับทฤษฎีออนโทโลยีที่มีการเชื่อมโยงความสัมพันธ์กันระหว่างข้อมูลความบกพร่องและสื่อการเรียนรู้ตลอดจนทรัพยากรการเรียนรู้ในด้านต่างๆเพื่อช่วยให้นักเรียนบกพร่องทางการเรียนรู้สามารถเข้าถึงสื่อการเรียนรู้ได้ตรงตามความต้องการของความบกพร่องแต่ละบุคคลมากขึ้นทำให้สามารถแก้ปัญหาการเรียนรู้ของผู้เรียนเอง รวมถึงครูผู้สอนหรือผู้ที่เกี่ยวข้องที่อาจมีความรู้ หรือ ไม่มีความรู้ในการเลือกสื่อการเรียนการสอนให้กับเด็กบกพร่องทางการเรียนรู้สามารถเลือกใช้สื่อการเรียนรู้ได้อย่างเหมาะสม

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

จากการศึกษาทฤษฎีและงานวิจัยทางด้านออนโทโลยีที่มีความเกี่ยวข้องเรื่องการทำคำแนะนำทรัพยากรการเรียนรู้ส่วนบุคคลสำหรับผู้เรียนบกพร่องทางการเรียนรู้สามารถสรุปได้ดังนี้

2.1 ความบกพร่องทางการเรียนรู้ (Specific Learning Difficulties)

เป็นความคิดปกติที่มีความยากลำบากในการฟัง การพูด การอ่าน การเขียน การให้เหตุผลรวมถึงความสามารถในการคิดคำนวณทางด้านคณิตศาสตร์[2] แบ่งเป็น 3 ประเภท ดังนี้

1) ความบกพร่องทางการอ่าน (Dyslexia) เป็นผู้เรียนที่มีปัญหาอ่านไม่ออกเขียนไม่ได้ จำตัวอักษรไม่ได้ อ่านไม่เป็นคำ

อ่านไม่เข้าใจหรือจับใจความไม่ได้ อ่านคำโดยสลับอักษร เช่น กบเป็นนก กาบเป็นบาก ความสามารถในการอ่านต่ำกว่าผู้เรียนคนอื่นๆซึ่งแตกต่างไปจากเด็กในระดับชั้นเดียวกัน

2) ความบกพร่องทางการเขียน (Dysgraphia) เกิดความผิดปกติทางการเขียน เช่น เขียนเป็นประโยคตามครูไม่ได้ ไม่สามารถลอกคำที่เขียนบนกระดานได้

3) ความบกพร่องทางคณิตศาสตร์ (Dyscalculia) เป็นเด็กที่มีปัญหาด้านสัญลักษณ์ต่างๆ ทางคณิตศาสตร์ โดยมีพฤติกรรมต่าง ๆ เช่น ไม่เข้าใจสัญลักษณ์ทางคณิตศาสตร์ว่าตัวดำเนินการ เช่น +, -, *, / เป็นค่าที่เพิ่มขึ้นหรือลดลง มีปัญหาในการจัดเรียงลำดับ

ประเภทของความบกพร่องรูปแบบต่าง ๆ ข้างต้นล้วนแต่จำเป็นต้องมีการใช้สื่อการเรียนรู้ลักษณะพิเศษที่สามารถช่วยแก้ปัญหาตามความบกพร่องแต่ละประเภทได้

2.2 เว็บเชิงความหมาย (Semantic Web)

เป็นส่วนขยายของ Web ปัจจุบันที่มีการอธิบายข้อมูลให้มีความหมายที่ชัดเจนมากขึ้น และช่วยให้คอมพิวเตอร์และคนสามารถเข้าใจข้อมูลในทิศทางเดียวกัน มีการกำหนดและเชื่อมโยงข้อมูลบนเว็บที่สามารถเพิ่มประสิทธิภาพในการสืบค้นได้ [3]

2.3 ออนโทโลยี (Ontology)

เป็นรูปแบบแนวความคิดที่แสดงถึงความรู้ซึ่งเป็นตัวแทนในการอธิบาย สร้างความสัมพันธ์ของวัตถุ (Objects) ที่บ่งบอกถึงลักษณะของคำศัพท์ที่จัดเก็บในฐานความรู้และนำเสนอเป็นองค์ความรู้ โดยการกำหนดชุดเงื่อนไขที่นำเสนอ ดังนี้ คลาส (Class) ความสัมพันธ์ (Relation) รูปแบบการทำงาน (Function) และวัตถุ (Object) ต่าง ๆ เป็นการอธิบายในสิ่งที่มนุษย์สามารถเข้าใจ มีหลักการที่เป็นแบบแผน การตีความหมาย ดังนั้นออนโทโลยีจึงเป็นคำสั่งของทฤษฎีทางตรรกะ [4]

Onintra และคณะทำการศึกษาและพัฒนาพัฒนาเทคโนโลยีพยากรณ์คำไทยที่เหมาะสมเพื่อช่วยในการเรียนรู้ของผู้เรียนบกพร่องทางการเรียนรู้[5] Sudhana และคณะศึกษารูปแบบออนโทโลยีที่สนับสนุนระบบการเรียนรู้อิเล็กทรอนิกส์ส่วนบุคคลโดยนำเสนอออนโทโลยีของผู้เรียนที่มีโดเมนคุณลักษณะที่สนับสนุนเพื่อจำลองการกำหนดรูปแบบการเรียนรู้โดยใช้โดเมนที่เฉพาะเจาะจงสำหรับการกำหนดค่าใน

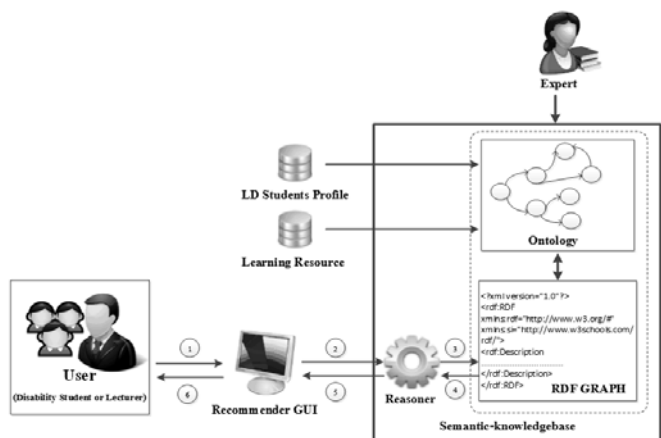
การสนับสนุนความรู้ในการเข้าถึงการเรียนรู้ส่วนบุคคลต่อการใช้สื่อการเรียนรู้ รวมถึงอธิบายประเภทของข้อมูลที่จะต้องพิจารณาในการพัฒนาออนโทโลยีที่ใช้สำหรับการจัดการสื่อการเรียนรู้ในระบบการเรียนรู้ออนไลน์ และมีการอธิบายถึงโครงสร้างสื่อการเรียนรู้ คือ ข้อมูลทั่วไป ข้อมูลเฉพาะ ข้อมูลการศึกษา ข้อมูลทางเทคนิคและข้อมูลสิทธิ [6] และ Nganji ได้นำเสนอสถาปัตยกรรมของการอนุมานความบกพร่องทางการเรียนรู้โดยใช้ออนโทโลยี [7]

2.4 การให้คำแนะนำบนพื้นฐานของความรู้

เป็นการให้คำแนะนำที่ชี้ให้เห็นถึงวัตถุประสงค์จากการวินิจฉัยจากความต้องการของผู้ใช้และการกำหนดค่าความต้องการ เทคนิคการให้คำแนะนำที่อธิบายประเภทของการอนุมานจากความรู้ที่เกี่ยวข้องของผู้ใช้เป็นรายการที่เฉพาะตรงตามความต้องการของผู้ใช้สามารถอนุมานเหตุผลจากความสัมพันธ์ระหว่างความต้องการและข้อเสนอแนะที่เป็นไปได้ รายละเอียดของผู้ใช้สามารถเป็นโครงสร้างของความรู้ที่สนับสนุนข้อสรุป [8]

3. การพัฒนาฐานความรู้

การพัฒนาฐานความรู้ซึ่งเป็นฐานความรู้เชิงความหมายจำเป็นต้องอาศัยโครงสร้างข้อมูลหรือออนโทโลยีในการพัฒนาฐานความรู้ นั้น มีกระบวนการพัฒนาดังต่อไปนี้



3.1 สถาปัตยกรรมระบบ

ภาพที่ 1: สถาปัตยกรรมของระบบการให้คำแนะนำทรัพยากรการเรียนรู้ส่วนบุคคลบนพื้นฐานของออนโทโลยีสำหรับผู้เรียนบกพร่องทางการเรียนรู้ (ontoLRforLD)

สถาปัตยกรรมของระบบการให้คำแนะนำทรัพยากรการเรียนรู้ส่วนบุคคลบนพื้นฐานของออนโทโลยีสำหรับผู้เรียนบกพร่องทางการเรียนรู้แสดงดังภาพที่ 1

1. อนุมานจาก Sparql Query เพื่อสอบถามจาก RDF GRAPH ซึ่งมีการพัฒนาจากออนโทโลยี
2. อนุมานแล้วได้ผลลัพธ์ที่ต้องการ เช่น สื่อการเรียนรู้ที่เหมาะสมกับนักเรียนบกพร่องทางการเรียนรู้ที่มีความบกพร่องทางด้านคณิตศาสตร์
3. ส่งผลลัพธ์ที่ได้จากการอนุมานไปยังส่วนติดต่อผู้ใช้
4. ผู้ใช้ได้รับคำแนะนำทรัพยากรการเรียนรู้ที่เหมาะสมที่แสดงในส่วนติดต่อผู้ใช้

3.2 การพัฒนาออนโทโลยี

ในงานวิจัยนี้ นำเสนอกระบวนการออกแบบและพัฒนาฐานความรู้บนพื้นฐานของออนโทโลยีตามวงจรการออกแบบออนโทโลยี (Ontology Life Cycle) ดังนี้ [9],[10]

1. ระบุโดเมนหรือขอบเขตของออนโทโลยี งานวิจัยนี้ต้องการที่จะพัฒนาองค์ความรู้ที่เกี่ยวข้องกับทรัพยากรการเรียนรู้และความต้องการใช้สื่อที่แตกต่างกันของผู้เรียนบกพร่องทางการเรียนรู้แต่ละประเภท พร้อมทั้งกำหนดคำถาม เช่น ทรัพยากรหรือสื่อการเรียนรู้ใดบ้างที่มีความเหมาะสมกับผู้เรียนบกพร่องทางการเรียนรู้ทางด้าน การอ่าน โดยผู้วิจัยประเมินความเหมาะสมของสื่อการเรียนรู้สำหรับนักเรียนบกพร่องทางการเรียนรู้แต่ละประเภท จึงได้ใช้ข้อมูลค่าความเหมาะสมเฉลี่ยจากผู้เชี่ยวชาญ ในการประเมินสื่อการเรียนรู้จำนวน 46 รายการ โดยกำหนดการเลือกใช้สื่อในระดับ 50% ขึ้นไป และแบ่งค่าความเหมาะสมออกเป็นสามช่วง คือ 50-69 :พอใช้(fair) 70-79:ดี(good) และ 80 ขึ้นไป:ดีมาก (excellent) ตัวอย่างการให้ค่าคะแนนความเหมาะสมของสื่อการเรียนรู้แสดงดังตารางที่ 1
2. พิจารณาออนโทโลยีที่มีอยู่กลับมาใช้ใหม่ ซึ่งผู้วิจัยทำการศึกษาจากงานวิจัยที่เกี่ยวข้อง โดยเฉพาะของ Nganji [7] ได้พัฒนาออนโทโลยีหลากหลายความบกพร่องด้วยกันแต่ยังไม่เจาะลึกด้านความบกพร่องทางการเรียนรู้จึงยังไม่เพียงพอที่จะนำกลับมาใช้ใหม่
3. ระบุประเด็นที่สำคัญของการพัฒนาออนโทโลยี โดยมีการกำหนดคำสำคัญที่ได้ทำการศึกษาจากองค์ความรู้ที่เกี่ยวข้อง

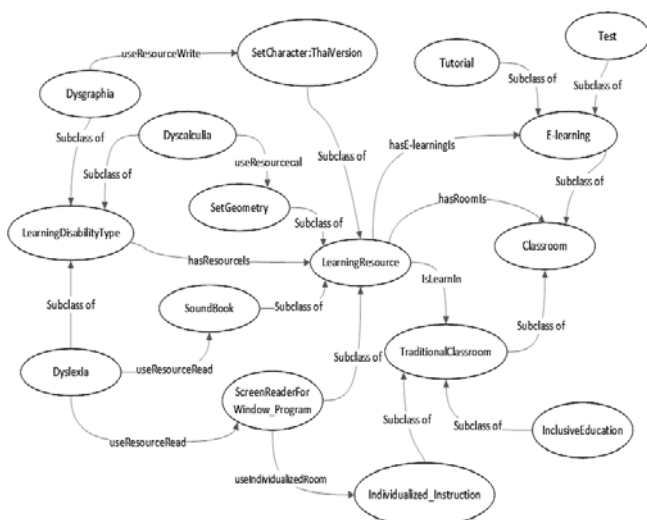
ตารางที่ 1: ตัวอย่างการประเมินค่าความเหมาะสมของสื่อการเรียนรู้โดยผู้เชี่ยวชาญ

ความบกพร่อง	เครื่องมือ/สื่อการเรียนรู้	ค่าเฉลี่ย (%)
ด้านการอ่าน	โปรแกรม IBM HPRReader	93.33
	โปรแกรม Scan and Read	93.33
	เครื่องจดบันทึกอักษรเบรลล์	26.67
ด้านการเขียน	ปากกาเน้นข้อความ แบบสะท้อนแสง	93.33
	กระดาษ A4 แบบมีบรรทัด	73.33
ด้านการคำนวณ	บทเรียนคอมพิวเตอร์ Math Flash	66.67
	เครื่องคิดเลขมีเสียงและตัวเลขขนาดใหญ่	80.00
	แท่งหลักนับเลข	73.33

และนำมานิยามเป็นคลาสต่างๆ เช่น ประเภทของความพิการทางการเรียนรู้ ความบกพร่องทางการอ่าน เป็นต้น

4. ระบุคลาสและลำดับชั้นของคลาส ผู้วิจัยนำองค์ความรู้ที่มีส่วนเกี่ยวข้องกับทรัพยากรการเรียนรู้สำหรับผู้เรียนบกพร่องทางการเรียนรู้กำหนดเป็นคลาสและลำดับชั้นของคลาส เช่น คลาส Dyslexia เป็นซับคลาสของ LearningDisabilityType ดังแสดงในภาพที่ 2

5. ระบุคุณลักษณะและความสัมพันธ์ให้กับคลาส โดยการสร้างความสัมพันธ์ระหว่างคลาสในขอบเขตที่สนใจ เช่น useResourceRead เป็นความสัมพันธ์ระหว่างคลาส Dyslexia และคลาส ScreenReaderForWindow_Program ดังแสดงในภาพที่ 2



ภาพที่ 2: ความสัมพันธ์แต่ละคลาส

6. กำหนดค่าของข้อมูลและสร้างเงื่อนไขต่างๆ เช่น ความสัมพันธ์ที่ชื่อ hasResourceIs มีโดเมน คือคลาส LearningDisabilityType และมีเรนจ์คือ LearningResource

7. การระบุค่าให้กับคลาส เป็นการกำหนดค่าข้อมูลต่างๆ ให้กับคลาส เช่น คลาส TraditionalClassroom เป็นอินสแตนซ์ของคลาส Classroom ดังตัวอย่างในภาพที่ 2

จากภาพที่ 2 สามารถสร้างเป็นออนโทโลยีในรูปแบบ OWL (Ontology Web Language) ที่เป็นภาษาสำหรับพัฒนาโครงสร้างออนโทโลยี สามารถอธิบายลำดับชั้นของคลาสและสร้างความสัมพันธ์ระหว่าง Resource ได้ดังตัวอย่างในภาพที่ 3

```

<!-- Object Properties -->
<?xml version="1.0"?>
...
<owl:ObjectProperty
rdf:about="http://www.semanticweb.org/LD#useIndividualizedRoom">
...
<rdfs:domain
rdf:resource="http://www.semanticweb.org/LD#ScreenReaderForWindow_Program"/>
</owl:ObjectProperty>

<!-- Classes -->
<owl:Class
rdf:about="http://www.semanticweb.org/LD#Dyslexia">
<rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/LD#LearningDisabilityType"/>
...
</rdf:RDF>
    
```

ภาพที่ 3: คลาสและสร้างความสัมพันธ์ระหว่างคลาส

จาก OWL ข้างต้นจะนำมาแปลงเป็นรูปแบบเอกสาร RDF เพื่อใช้ในกระบวนการอนุมานคำตอบดังตัวอย่างในภาพที่ 4

```

<Id:disability
rdf:about="http://www.semanticweb.org/LD#Dyslexia">
<Id:useResourceRead
rdf:resource="http://www.semanticweb.org/LD#ScreenReaderForWindow_Program"/>
...
rdf:resource="http://www.semanticweb.org/LD#SoundBook"/>
</Id:useResourceRead>
</Id:disability>
    
```

ภาพที่ 4: ตัวอย่างเอกสารเอกสาร RDF

3.3 การอนุมานคำตอบ

การทดสอบแนวคิดว่าระบบสามารถค้นคืนทรัพยากรการเรียนรู้ตามความต้องการของผู้เรียนบกร่องทางการเรียนรู้ได้ถูกต้องและครบถ้วนหรือไม่ ผู้วิจัยใช้ภาษา Sparql Query ในการสอบถามโดยมีการค้นคืนจากเอกสาร RDF ผ่านทางโปรแกรมประยุกต์ที่พัฒนาด้วยภาษา PHP นั้น เช่น ต้องการสืบค้นทรัพยากรการเรียนรู้ของผู้เรียนที่มีความบกพร่องทางการอ่าน โดยต้องการทราบข้อมูลดังนี้ คือ ประเภทของความบกพร่อง ทรัพยากรการเรียนรู้ต่างๆ เช่น รายชื่อสื่อการเรียนรู้ ใช้รูปแบบห้องเรียนแบบใด สมควรที่จะใช้รูปแบบการเรียนแบบอิเล็กทรอนิกส์อย่างไร รวมไปถึงระดับความเหมาะสมของสื่อการเรียนรู้ที่จะนำไปใช้งานที่ผ่านการตรวจสอบโดยผู้เชี่ยวชาญ ดังแสดงตัวอย่างในภาพที่ 5

```
SELECT ?LDType , ?LearningResource ,?Satisfaction ,?class , ?E_learningType
WHERE
    (?a ld:LDTypeName ?LDType)
    (?k ld:useResourceWrite ?d)
    (?d ld:LearningResourceName
?LearningResource)
    (?d ld:Satisfaction ?Satisfaction )
    (?f ld:TraditonalClassName ?class)
    (?v ld:E-learningTypeName ?E_learningType)

    AND ?LDType=="บกพร่องทางการอ่าน"

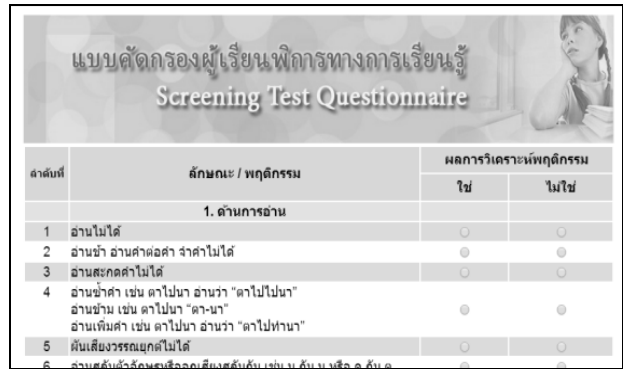
USING rdf FOR <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
    Id FOR http://www.semanticweb.org/LD#
```

ภาพที่ 5: Sparql Query การค้นคืนทรัพยากรการเรียนรู้สำหรับ ผู้เรียนบกพร่องทางการอ่าน

4 ผลการดำเนินงานและผลการประเมิน

4.1 ส่วนการติดต่อผู้ใช้

ให้คำแนะนำทรัพยากรการเรียนรู้มีส่วนติดต่อผู้ใช้ที่พัฒนาด้วยภาษา PHP ติดต่อกับฐานความรู้ออนโทโลยีที่ออกแบบไว้ ภาพที่ 6 แสดงส่วนการคัดกรอง (Screening Test) และภาพที่ 7 แสดงผลลัพธ์จากการอนุมานคำตอบของระบบ ซึ่งประกอบไปด้วยข้อมูล ชนิดข้อมูลความบกพร่อง (?LDType) รายละเอียดสื่อที่ได้รับการแนะนำ ถึงแวดล้อมในการเรียน ระดับความเหมาะสม(? Satisfaction) เป็นต้น



ภาพที่ 6: หน้าคัดกรองความบกพร่องทางการเรียนรู้

No.	?LDType	?LearningResource	? Satisfaction	?class	?E_learningType
1	Liberal ศึกษาทางด้าน การอ่าน	Liberal ฝึกอ่าน Kurzweil 3000	Liberal ดี มาก	Liberal ศึกษาระดับชั้น	Liberal แบบฝึกหัดและฝึกปฏิบัติ
2	Liberal ศึกษาทางด้าน การอ่าน	Liberal ฝึกอ่าน Scan and Read	Liberal ดี มาก	Liberal ศึกษาระดับชั้น	Liberal แบบฝึกหัดและฝึกปฏิบัติ
3	Liberal ศึกษาทางด้าน การอ่าน	Liberal เครื่องมือฟังหนังสือเสียง สำหรับหนังสือระบบเคซี	Liberal พอใช้	Liberal ศึกษาระดับชั้น	Liberal แบบฝึกหัดและฝึกปฏิบัติ
4	Liberal ศึกษาทางด้าน การอ่าน	Liberal เครื่องเล่น MP4 ฟังหนังสือได้	Liberal ดี มาก	Liberal ศึกษาระดับชั้น	Liberal แบบฝึกหัดและฝึกปฏิบัติ
5	Liberal ศึกษาทางด้าน การอ่าน	Liberal เครื่องเล่น DVD	Liberal ดี	Liberal ศึกษาระดับชั้น	Liberal แบบฝึกหัดและฝึกปฏิบัติ

ภาพที่ 7: ผลการค้นคืนสารสนเทศของระบบจากออนโทโลยี

4.2 การทดสอบฐานความรู้ออนโทโลยี

ผู้วิจัยได้ทดสอบเพื่อการพิจารณาออนโทโลยีที่พัฒนานั้นสามารถตอบคำถามได้จริงหรือไม่โดยใช้การประเมินผลโดยผู้เชี่ยวชาญใน 5 ประเด็นดังต่อไปนี้ กำหนดขอบเขตของเรื่องที่ชัดเจน (Scope:S) ความถูกต้องของข้อมูล (Accuracy:A) ความเข้าใจข้อมูลที่ต้องการสื่อ (Understand:U) มีเนื้อหาที่ครอบคลุมตามวัตถุประสงค์ (Comprehensive:C) การนำไปใช้และประโยชน์ (Benefit:B) มีระดับคะแนนตั้งแต่ 1-5 คะแนน (เห็นด้วยน้อยสุดไปมากที่สุด) โดยทำการหาผลการทดสอบโดยรวมจาก 5 ประเด็นข้างต้น ดังสมการที่ 1 [10]

$$\Sigma \text{ score} = S(w_s) + A(w_a) + U(w_u) + C(w_c) + B(w_b) \quad (1)$$

กำหนดให้ค่าถ่วงน้ำหนัก Weight (w) = 0.20 โดย $\Sigma w = 1$ จากการประเมินฐานความรู้ออนโทโลยีทรัพยากรการเรียนรู้สำหรับผู้เรียนบกพร่องทางการเรียนรู้โดยผู้เชี่ยวชาญทั้ง 5 ประเด็นผลที่ได้อยู่ในเกณฑ์ที่ดี คือ ค่าเฉลี่ย (\bar{x}) = 4.16 ส่วนเบี่ยงเบน (S.D) = 0.11 และ ค่า $\Sigma \text{ score} = 4.16$ ดังตารางที่ 2

ตารางที่ 2: ผลการทดสอบฐานความรู้ออนโทโลยีทรัพยากรการเรียนรู้ สำหรับผู้เรียนบทพร้อมทางการเรียนรู้

ประเด็นในการทดสอบ	\bar{x}	S.D	score
กำหนดขอบเขตของเรื่องที่ชัดเจน(S)	4.4	0.54	$W_s = 0.88$
ความถูกต้องของข้อมูล(A)	4.2	0.44	$W_A = 0.84$
ความเข้าใจข้อมูลที่ต้องการสื่อ(U)	4	0.70	$W_U = 0.80$
มีเนื้อหาที่ครอบคลุม(C)	3.8	0.44	$W_C = 0.76$
การนำไปใช้และประโยชน์(B)	4.4	0.54	$W_B = 0.88$
Total	4.16	0.11	$\Sigma \text{ score} = 4.16$

5. สรุป

ในงานวิจัยนี้ผู้วิจัยได้รวบรวมองค์ความรู้ที่เกี่ยวข้องเช่น ความบกพร่องทางการเรียนรู้ในด้านต่างๆ สื่อการเรียนรู้ที่เหมาะสมกับแต่ละประเภทของความบกพร่องทางการเรียนรู้ โดยมีผู้เชี่ยวชาญตรวจสอบและกำหนดระดับความพึงพอใจของสื่อรวมถึงทรัพยากรต่างๆที่เกี่ยวข้องมาสร้างความสัมพันธ์เชื่อมโยงกัน พัฒนาเป็นระบบให้คำแนะนำทรัพยากรการเรียนรู้โดยอัตโนมัติบนฐานความรู้ออนโทโลยี สามารถตอบสนองความต้องการส่วนบุคคลของแต่ละประเภทความบกพร่องทางการเรียนรู้โดยจากการประเมินผลการทดสอบฐานความรู้ออนโทโลยีทรัพยากรการเรียนรู้สำหรับผู้เรียนบกพร่องทางการเรียนรู้คือค่าเฉลี่ย (\bar{x}) = 4.16 ส่วนเบี่ยงเบน (S.D) = 0.11 และ ค่า $\Sigma \text{ score} = 4.16$ ถือว่าผลการทดสอบอยู่ในเกณฑ์ที่ดี ซึ่งผลการวิจัยนี้สามารถนำองค์ความรู้จากการพัฒนาออนโทโลยีไปประยุกต์ใช้กับงานต่างๆได้ในอนาคต และอาจขยายการเชื่อมโยงไปยังผู้เชี่ยวชาญด้านสื่อการเรียนรู้สำหรับเด็กบกพร่องในการจัดการฐานความรู้ และขยายฐานออนโทโลยีให้มีความครอบคลุม ทันสมัยรองรับกับสื่อใหม่ๆ ที่เกิดขึ้นในอนาคต

6. กิตติกรรมประกาศ

งานวิจัยที่ได้รายงานในบทความนี้ได้รับการสนับสนุนจากหน่วยวิจัยการจัดการความรู้และเทคโนโลยีเว็บเชิงความหมาย สถาบันวิจัยและพัฒนา มหาวิทยาลัยทักษิณ

เอกสารอ้างอิง

- [1] J. T. Nganji, (2012). "Designing Disability-Aware E-Learning Systems : Disabled Students' Recommendations," *International Journal of Advanced Science and Technology*. Vol.48, 61-70.
- [2] ผดุง อารยะวิญญู. (2544). *เด็กที่มีปัญหาในการเรียนรู้ Learning Disabilities* (พิมพ์ครั้งที่ 2). กรุงเทพฯ : แวนแก้ว.
- [3] T. Berners-Lee, J. Hendler and O. Lassila, "The Semantic Web", *Scientific American*, May 2001, p. 29-37.
- [4] T. R. Gruber. (1993). "Toward Principles for the Design of Ontologies Used for Knowledge Sharing," *International Workshop on Formal Ontology*. Vol.43(5-6), 907-928. Retrieved June 2,2013, from <http://www.sciencedirect.com/science/article/pii/S1071581985710816>
- [5] O. Poobrasert and A. Wongteeratana, (2012). "Assistive Technology : Design & Development for Children with Learning Disabilities," *IEEE Technology and Engineering Education*. Vol.7(4), 38-43.
- [6] Madhu Sudhana (2013). "An Ontological Approach to Support Personalized E-Learning System," *International Journal of Computer Applications*. Vol.61(2).
- [7] J. T. Nganji, M. Brayshaw and B. Topsett, (2011) *Ontology-based E-learning Personalisation for Disabled Students in Higher Education*. [online]. from : <http://www.heacademy.ac.uk>.
- [8] R. Burke, (2002). "Hybrid Recommender Systems: Survey and Experiments," *User Modeling and User-Adapted Interaction*. Vol.12(4), 331-370.
- [9] N.F. Noy and D.L.Mc.Guinness, *Ontology Development 101 : A Guide to Creating Your First Ontology* [Online]. Available from:<http://ksl.stanford.edu/people/dlm/papers/ontology-tutorial-noy-mcguinness-abstract.html> [10 January, 2011].
- [10] วิชดา โชติรัตน์. 2011. "การพัฒนาฐานความรู้ออนโทโลยีสำหรับวิเคราะห์ข่าวออนไลน์โดยอัตโนมัติ." *Information Technology Journal*, Vol. 7, No.14. [วารสารออนไลน์] เข้าถึงจาก: <http://www.it.kmutnb.ac.th/journal/pdf/vol14/ch03.pdf>

ขั้นตอนวิธีการทำนายค่าคะแนนความชอบที่มีประสิทธิภาพสำหรับระบบแนะนำเพลง โดยใช้วิธีการเทรนด์เบสร่วมกับข้อมูลไมโครโพรไฟล์

The Efficient Rating Prediction Algorithm for Music Recommender System by Using Tendencies Based Algorithm with Micro Profiles

สุเมธ ดาราพิสูทธิ์ (Sumet Darapisut)¹ และ จักริน สุขสวัสดิ์ชน (Jakkarin Suksawatchon)²

คณะวิทยาการสารสนเทศ มหาวิทยาลัยบูรพา

¹dearsumet@gmail.com, ²jakkaman@gmail.com

บทคัดย่อ

งานวิจัยนี้นำเสนอขั้นตอนวิธีการทำนายค่าคะแนนความชอบที่มีประสิทธิภาพสำหรับระบบแนะนำเพลงชื่อว่า “Tendencies Based Algorithm with Micro Profiles & Sparsity Problem (TBMS)” ซึ่งเป็นการประยุกต์ใช้วิธีการ Tendencies Based Algorithm ร่วมกับบริบททางด้านเวลาแนะนำเพลง โดยการแบ่งข้อมูลการฟังเพลงของผู้ใช้ออกเป็นข้อมูลการฟังเพลงในแต่ละช่วงเวลาเรียกว่า “Micro Profiles” ทั้งหมด 5 ชุดด้วยกันรวมถึงงานวิจัยนี้ยังพิจารณาการแก้ไขปัญหาความเบาบางของข้อมูล โดยการเติมค่าข้อมูลก่อนการทำนายเพื่อเพิ่มประสิทธิภาพของการทำนายที่ดีขึ้นกว่าขั้นตอนการทำนายแบบเดิมที่นิยมใช้วิธีการกรองร่วม ส่วนข้อมูลที่ใช้ในงานวิจัยนี้เป็นข้อมูลการฟังเพลงจากฐานข้อมูลของ last.fm ที่ประกอบด้วยประวัติการฟังเพลงของผู้ใช้แต่ละคนซึ่งจากการทดลองแสดงให้เห็นว่าวิธีการที่นำเสนอ TBMS นี้มีประสิทธิภาพ โดยให้ผลการทำนายความชอบที่แม่นยำและมีความรวดเร็วในการประมวลผลมากกว่าวิธีการกรองร่วมแบบดั้งเดิม

คำสำคัญ: ระบบแนะนำเพลง เวลา ไมโครโพรไฟล์ ปัญหาความเบาบางของข้อมูล

Abstract

This research presents the efficient rating prediction algorithm for music recommender system named “Tendencies Based Algorithm with Micro Profiles & Sparsity Problem (TBMS)”, which adopts Tendencies

Based Algorithm incorporated with time contextual. The user profile is splited into several sub-profile based on time contextual called “Micro Profiles”; totally 5 sets, each representing the user in particular context. In addition, the research also considers to solve the sparsity problem by imputation the missing values to enhance the performance of prediction process better than the traditional algorithm -- collaborative filtering technique. For all experiments, this research used last.fm data set which is composed of the music listening of each users. From our experiments, we have found that the proposed method TBMS is effective. It is give more accuracy and much more efficient computationally than tradition collaborative filtering.

Keyword: music recommender system, time, Micro Profiles, sparsity problem.

1. บทนำ

ระบบแนะนำเพลงหรือที่เรียกว่าโปรแกรมแนะนำเพลง เป็นโปรแกรมที่ช่วยในการค้นหาและแนะนำเพลงที่ผู้ใช้งานต้องการฟัง โดยจะวิเคราะห์จากประวัติการฟังเพลงของผู้ใช้ว่าชอบฟังเพลงแนวใด และโปรแกรมก็จะสร้างรายการแนะนำเพลงสำหรับผู้ใช้นั้น เช่นแนะนำตามความชอบ แนะนำตามแนวเพลง แนะนำตามศิลปิน หรือแนะนำตามช่วงเวลาที่ยังฟังเพลง โปรแกรมแนะนำเพลงที่เป็นที่นิยมในปัจจุบันได้แก่ Pandora, Songza, Last.fm, iTunes Genius เป็นต้น

ในปัจจุบันเทคนิคที่ใช้สร้างระบบแนะนำเพลงที่นิยมใช้ ได้แก่ วิธีการกรองร่วม (Collaborative filtering) [1,2,3] การแนะนำจะพิจารณาแนะนำเพลงให้กับผู้ใช้จากเพลงที่บุคคลอื่นมีความชอบคล้ายกับผู้ใช้ โดยที่ไม่จำเป็นต้องรู้จักคนเหล่านั้น แล้วนำการแนะนำของบุคคลเหล่านั้นมาช่วยตัดสินใจเลือกเพลง แต่วิธีการนี้ใช้ค่าความชอบหรือคะแนน (Rating) ที่ผู้ใช้มีต่อเพลงแต่ละเพลงซึ่งต้องจัดเก็บค่าไว้ให้กับระบบแนะนำเสียก่อน และหลายครั้งทำให้วิธีการกรองร่วมมีปัญหาที่เกิดขึ้นสำหรับการแนะนำเช่น ปัญหาความเบาบางของข้อมูล (Sparsity Problem) และการแนะนำเพลงที่ไม่ค่อยเป็นที่รู้จัก ซึ่งเกิดจากสาเหตุที่ระบบแนะนำเพลงโดยทั่วไปมีการใช้ปัจจัยเพียงแค่ 2 ปัจจัยสำหรับสร้างระบบแนะนำคือ เพลงและผู้ใช้ ซึ่งไม่เพียงพอต่อการสร้างระบบแนะนำให้มีประสิทธิภาพ แต่ถ้านำข้อมูลบริบทของผู้ใช้ (Contextual information) มารวมในการพิจารณาเช่น แนวเพลงที่ผู้ใช้ชอบฟัง รวมไปถึงช่วงเวลาที่แตกต่างกันผู้ใช้อาจจะชอบฟังแนวเพลงที่ไม่เหมือนกัน ก็จะสามารถพิจารณาแนะนำเพลงได้ตรงกับความต้องการของผู้ใช้มากขึ้น

งานวิจัย [4] ได้ศึกษาและวิเคราะห์พฤติกรรมของผู้ฟังเพลงซึ่งพบว่าบริบททางด้านเวลาของการฟังเพลงมีความแตกต่างจากบริบทอื่นๆ เพราะบางแนวเพลงมีการฟังเฉพาะบางช่วงเวลาของวัน เช่นเพลงแนวแดนซ์ป๊อป, บัลลาดและร็อกถูกฟังมากในตอนกลางวันและฟังน้อยลงในตอนเย็น ส่วนงานวิจัย [5,6,7] มีการนำเสนอระบบแนะนำเพลงโดยพิจารณามิติทางด้านเวลา โดยทำการแบ่งข้อมูลของผู้ใช้ตามช่วงเวลา ที่เรียกว่า “Micro Profiles” เช่น ช่วงเวลาของวัน วันหยุดหรือวันทำงาน เป็นต้น และใช้ข้อมูล Micro Profiles ร่วมกับเทคนิคการกรองร่วม ซึ่งผลการวิจัยพบว่าการใช้บริบทของเวลาร่วมกับเทคนิคการกรองร่วมจะให้ความถูกต้องในการแนะนำสูงกว่าการใช้การกรองร่วมเพียงอย่างเดียว แต่อย่างไรก็ตามงานวิจัยดังกล่าวก็ยังพบปัญหาที่เกิดจากความเบาบางของข้อมูล และดำเนินการกับข้อมูลขนาดใหญ่ ซึ่งทำให้การแนะนำเพลงมีประสิทธิภาพการทำงานลดลง

งานวิจัย [8] ได้นำเสนอวิธีการสำหรับระบบแนะนำข้อมูลสำหรับการกรองร่วมที่ชื่อว่า “Tendencies Based Algorithm” ซึ่งจะดูแนวโน้มของความชอบผู้ใช้และสินค้า ข้อดีของวิธีการนี้

คือ เวลาในการคำนวณที่น้อยลงมากเมื่อเทียบกับระบบแนะนำแบบการกรองร่วมและยังสามารถใช้กับข้อมูลที่มีปริมาณมากได้อย่างไรก็ตามปัญหาความเบาบางของข้อมูลก็ยังเป็นจุดด้อยวิธีการนี้ และในปัจจุบันยังไม่มียานวิจัยที่เกี่ยวกับระบบแนะนำเพลงแบบกรองร่วมโดยนำวิธีการ Tendencies Based มาใช้เพื่อทดสอบประสิทธิภาพการแนะนำเพลงโดยพิจารณาพร้อมกับบริบทของเวลา

ดังนั้นงานวิจัยนี้จะนำเสนอวิธีการทำนายค่าคะแนนที่มีประสิทธิภาพสำหรับระบบแนะนำเพลงชื่อว่า “Tendencies Based Algorithm with Micro Profiles & Sparsity Problem” ซึ่งจะประยุกต์ใช้วิธีการ Tendencies Based Algorithm ร่วมกับการใช้บริบททางด้านเวลาแนะนำเพลง รวมถึงการแก้ไขปัญหาความเบาบางของข้อมูล โดยการเติมค่าข้อมูลก่อนการทำนายซึ่งวิธีการที่นำเสนอจะช่วยลดเวลาที่ใช้ในการคำนวณสำหรับระบบแนะนำที่ใช้วิธีการกรองร่วมแบบเดิม และเพิ่มความถูกต้องให้กับระบบแนะนำเพลงตามช่วงเวลาให้กับผู้ใช้

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ระบบแนะนำ (Recommender System)

เทคนิคที่นิยมในการพัฒนาระบบแนะนำคือวิธีการกรองร่วม ซึ่งแบ่งออกเป็น 2 ลักษณะ กล่าวคือพิจารณาจากตัวผู้ใช้เป็นสำคัญ (User based) และพิจารณาจากไอเท็มเป็นสำคัญ (Item based) [9] แต่ในงานวิจัยนี้เน้นวิธีการกรองร่วมที่พิจารณาจากผู้ใช้เป็นสำคัญ ดังนั้นจึงขอกว่าเฉพาะวิธีการนี้เท่านั้น โดยขั้นตอนวิธีการกรองร่วมที่พิจารณาจากผู้ใช้เป็นสำคัญนั้น เริ่มต้นจากการหาความคล้ายคลึงระหว่างผู้ใช้ u กับผู้ใช้ v โดยวิธีการหาค่าความคล้ายคลึงที่นิยม [3] ได้แก่

วิธีการหาค่าความคล้ายคลึงแบบโคไซน์ (Cosine-based Similarity)

$$sim(u, v) = \frac{\sum_i r_{u,i} r_{v,i}}{\sqrt{\sum_i r_{u,i}^2} \sqrt{\sum_i r_{v,i}^2}} \quad (1)$$

วิธีการหาค่าความคล้ายคลึงแบบโคไซน์แบบปรับแก้ (Adjusted Cosine Similarity)

$$sim(u, v) = \frac{\sum_i (r_{u,i} - \bar{r}_i)(r_{v,i} - \bar{r}_i)}{\sqrt{\sum_i (r_{u,i} - \bar{r}_i)^2} \sqrt{\sum_i (r_{v,i} - \bar{r}_i)^2}} \quad (2)$$

การหาค่าความคล้ายคลึงแบบค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Pearson correlation coefficient)

$$sim(u, v) = \frac{\sum_{i \in I_u \cap I_v} (r_{u,i} - \bar{r}_u)(r_{v,i} - \bar{r}_v)}{\sqrt{\sum_{i \in I_u \cap I_v} (r_{u,i} - \bar{r}_u)^2} \sqrt{\sum_{i \in I_u \cap I_v} (r_{v,i} - \bar{r}_v)^2}} \quad (3)$$

และเมื่อหาค่าความคล้ายโดยสมการ(1)-(3) แล้วก็จะเข้าสู่ขั้นตอนการทำนายว่าผู้ใช้แต่ละคนจะมีค่าคะแนนความชอบกับแต่ละไอเท็มเท่าไร โดยใช้สมการ (4)

$$p_{u,i} = \bar{r}_u + \frac{\sum_{v \in N} sim(u, v)[r_{v,i} - \bar{r}_v]}{\sum_{v \in N} |sim(u, v)|} \quad (4)$$

โดย $r_{u,i}$ คือคะแนนความชอบที่ผู้ใช้ u ให้อิเท็ม i
 \bar{r}_u คือค่าเฉลี่ยของผู้ใช้ u
 \bar{r}_v คือค่าเฉลี่ยของผู้ใช้ v
 $i \in I_u \cap I_v$ คือไอเท็มที่ผู้ใช้ทั้งสองให้คะแนนร่วมกัน
 \bar{r}_i คือค่าเฉลี่ยของไอเท็มที่ i
 N คือจำนวนเพื่อนบ้านของผู้ใช้ u

2.2 Tendencies Based Algorithm [8]

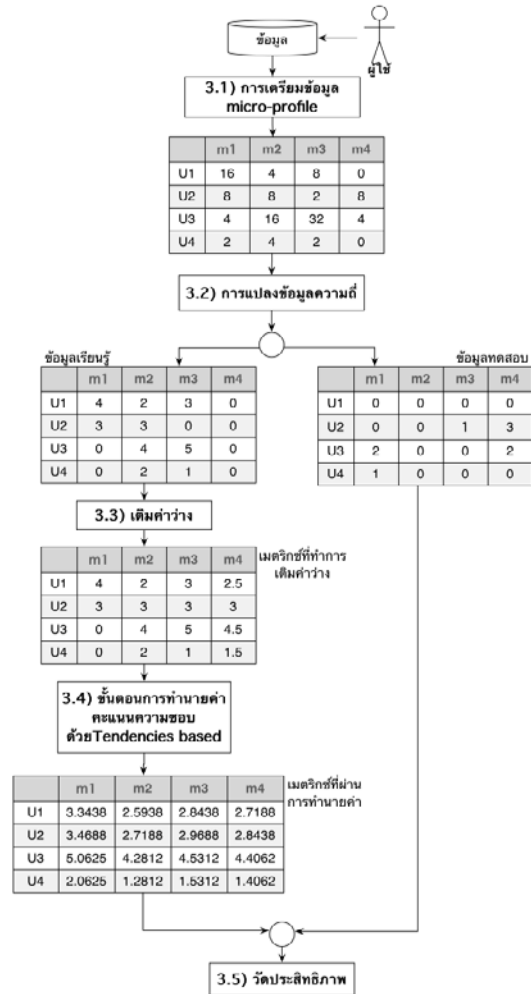
วิธีการนี้เป็นวิธีการกรองร่วมอีกเทคนิคหนึ่งที่ทำนายค่าความชอบของผู้ใช้ที่มีต่อไอเท็มจากค่าแนวโน้มความชอบของผู้ใช้กับไอเท็มนั้น โดยจะพิจารณาจากความแตกต่างระหว่างคะแนนของไอเท็มและค่าเฉลี่ยของไอเท็มหรือค่าเฉลี่ยของผู้ใช้ซึ่งทำให้ใช้เวลาการคำนวณที่รวดเร็วและทำงานได้ดีกับข้อมูลที่มีความหนาแน่นต่ำ

3. วิธีการที่นำเสนอ

งานวิจัยนี้ได้นำเสนอวิธีการใหม่ที่เรียกว่า “Tendencies Based Algorithm with Micro Profiles & Sparsity Problem (TBMS)” ซึ่งมีภาพรวมการทำงานดังภาพที่ 1

3.1 การเตรียมข้อมูล Micro Profiles

ในงานวิจัยนี้ใช้ข้อมูลการฟังเพลงจากฐานข้อมูลของ last.fm¹ [10] ประกอบด้วยประวัติการฟังเพลงของผู้ใช้แต่ละคนซึ่งจะต้องนำมาแปลงข้อมูลให้อยู่ในรูปแบบของความถี่ในการ



ภาพที่ 1: ขั้นตอนการทำงานของวิธีการที่นำเสนอ

ฟังเพลงแต่ละเพลงของผู้ใช้แต่ละคนจากนั้นนำข้อมูลความถี่ทั้งหมดแบ่งออกเป็น 1) ข้อมูลทั้งหมดที่ไม่ได้แบ่งตามช่วงเวลา และ 2) ข้อมูลทั้งหมดที่แบ่งออกเป็น Micro Profiles ตามช่วงเวลา จำนวน 5 ชุดด้วยกันคือช่วงเช้า, ช่วงกลางวัน, ช่วงเย็น วันหยุดและวันธรรมดา ซึ่งจะมีข้อมูลที่ใช้ในการทดลองทั้งหมด 6 ชุดดังแสดงในตารางที่ 1

ตารางที่ 1: ข้อมูลที่ใช้ในการทดลอง

ข้อมูล	จำนวนผู้ใช้	จำนวนเพลง
ข้อมูลทั้งหมด	222	183,322
ตอนเช้า (00:00-11:59 น.)	182	130,519
ตอนกลางวัน(12:00-17:59 น.)	184	79,231
ตอนเย็น (18:00-23:59 น.)	187	103,393
วันหยุด (เสาร์-อาทิตย์)	212	98,387
วันธรรมดา(วันจันทร์-วันศุกร์)	190	162,442

¹ <http://www.lastfm.com>

3.2 การแปลงข้อมูลความถี่

เนื่องจากข้อมูลความถี่ที่ได้มาจากขั้นตอนที่ 3.1 มีความแตกต่างของข้อมูลค่อนข้างมาก คือบางเพลงถูกฟังเพียงครั้งเดียว แต่บางเพลงถูกฟังหลายพันครั้ง ดังนั้นในงานวิจัยนี้จึงได้ทำการแปลงความถี่ของการฟังเพลงออกเป็นคะแนนความนิยมโดยใช้วิธีการ log scale[11] ดังสมการ (5)

$$\widehat{r_{u_i, m_j}} = \log_2(r_{u_i, m_j} + 1) \quad (5)$$

โดยที่ r_{u_i, m_j} คือความถี่ในการฟังเพลงที่ m_j ของผู้ใช้ u_i

3.3 การเติมค่าว่าง

เนื่องจากผู้ใช้ไม่ได้ฟังเพลงทุกเพลงที่มีในฐานข้อมูล ดังนั้นจึงทำให้เพลงที่ผู้ใช้คนนั้นไม่ได้ฟัง จึงไม่มีข้อมูลหรือเกิดค่าว่างเกิดขึ้น ซึ่งทำให้ข้อมูลมีความหนาแน่นน้อยและส่งผลต่อความถูกต้องของระบบแนะนำเพลง และเกิดปัญหาที่เรียกว่า *ปัญหาความเบาบางของข้อมูล* เราจะใช้วิธีการเติมค่าว่างด้วยการเติมคะแนนความชอบเฉลี่ยของแนวเพลงนั้น

ตารางที่ 2: ตัวอย่างเมทริกซ์ที่เก็บค่าคะแนนความชอบที่ผู้ใช้แต่ละคนฟังเพลง โดยที่ ? คือเพลงที่ผู้ใช้ไม่เคยฟัง

	เพลงที่ 1	เพลงที่ 2	เพลงที่ 3	เพลงที่ 4
	jazz	pop	pop	pop
ผู้ใช้ u_1	4	2	3	?
ผู้ใช้ u_2	3	3	?	?
ผู้ใช้ u_3	?	4	5	?
ผู้ใช้ u_4	?	2	1	?

ตัวอย่างในตารางที่ 2 แสดงข้อมูลค่าคะแนนความชอบที่ผู้ใช้แต่ละคนชอบฟังเพลงโดยที่ ? คือเพลงที่ผู้ใช้ไม่เคยฟังถ้าต้องการเติมค่าว่าง (?) ของผู้ใช้ u_1 ในเพลงที่ 4 ซึ่งเป็นเพลงแนวป๊อป วิธีการการคำนวณค่าคะแนนความชอบเฉลี่ยจะนำค่าคะแนนความชอบเพลงที่ 2 และ เพลงที่ 3 ซึ่งเป็นเพลงแนวป๊อปเช่นเดียวกันกับเพลงที่ 4 มาคำนวณจะได้ $(2+3)/2 = 2.5$ ดังแสดงในตารางที่ 3

ตารางที่ 3: ตัวอย่างเมทริกซ์ที่เก็บค่าคะแนนความชอบที่มีการเติมค่าว่าง

	เพลงที่ 1	เพลงที่ 2	เพลงที่ 3	เพลงที่ 4
	jazz	pop	pop	pop
ผู้ใช้ u_1	4	2	3	2.5
ผู้ใช้ u_2	3	3	3	3
ผู้ใช้ u_3	?	4	5	4.5
ผู้ใช้ u_4	?	2	1	1.5

3.4 ขั้นตอนการทำนายค่าคะแนนความชอบด้วยวิธีการ

Tendencies Based Algorithm

ขั้นตอนนี้เป็นการคำนวณค่าคะแนนความชอบของผู้ใช้แต่ละคน โดยจะพิจารณาจากแนวโน้มในการชอบฟังเพลงของผู้ใช้ ซึ่งประยุกต์ใช้วิธีการ Tendencies Based Algorithm มาใช้ในการทำนาย โดยกำหนดค่าตัวแปรดังตารางที่ 4

ตารางที่ 4: นิยามค่าตัวแปร

สัญลักษณ์	ความหมาย
$U = \{u_1, u_2, \dots, u_i, \dots, u_m\}$	เซตของผู้ใช้
$M = \{m_1, m_2, \dots, m_j, \dots, m_n\}$	เซตของเพลง
V	เมทริกซ์ของค่าคะแนนความชอบขนาด $m \times n$
v_{u_i, m_j}	คะแนนความชอบของเพลง m_j ที่ผู้ใช้ u_i ให้ไว้
$V_{u_i} = \{v_{u_i, m_j} \in V m_j \in M\}$	เซตของค่าคะแนนความชอบของเพลงที่ผู้ใช้ u_i
$V_{\cdot, m_j} = \{v_{u_i, m_j} \in V u_i \in U\}$	เซตค่าคะแนนความชอบที่ผู้ใช้ให้กับเพลง m_j
\bar{V}_{u_i}	ค่าเฉลี่ยของผู้ใช้ u_i
\bar{V}_{\cdot, m_j}	ค่าเฉลี่ยของเพลง m_j

ขั้นตอนวิธีการการทำนายค่าคะแนนความชอบดังนี้

ขั้นที่ 1 หาค่าแนวโน้มความชอบของผู้ใช้ u_i และเพลง m_j ตามสมการ (6) และสมการ (7)

$$u_{b_{u_i}} = \frac{\sum_{m_j \in M_{u_i}} (v_{u_i, m_j} - \bar{v}_{\cdot, m_j})}{|M_{u_i}|} \quad (6)$$

$$i_{b_{m_j}} = \frac{\sum_{u_i \in U_i} (v_{u_i, m_j} - \bar{v}_{u_i})}{|U_i|} \quad (7)$$

โดยที่ M_{u_i} แทนจำนวนเพลงที่ผู้ใช้ u_i ฟัง

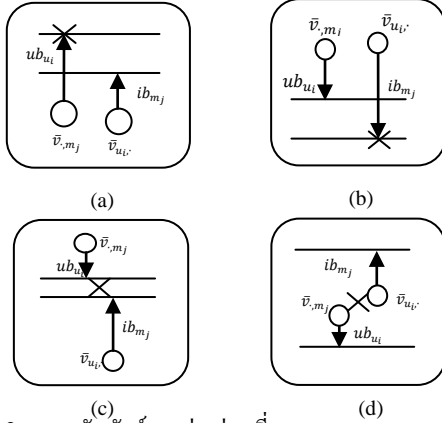
U_i แทนจำนวนผู้ใช้ที่ฟังเพลง m_j

ขั้นที่ 2 ทำนายค่าคะแนนความพึงพอใจของผู้ใช้ u_i กับเพลง m_j จะแบ่งการพิจารณาออกเป็น 4 กรณีตามหลักการของ Tendencies Based ดังภาพที่ 2 โดยแบ่งการทำนายออกเป็น 4 กรณีดังนี้

กรณีที่ 1 ภาพที่ 2(a) ทำนายจากแนวโน้มความชอบที่เป็นบวกทั้งผู้ใช้และเพลงหมายความว่าผู้ใช้มีแนวโน้มที่จะแนะนำ

เพลงที่ฟังมีค่าสูงในหลายเพลง โดยถ้าค่าแนวโน้มของผู้ใช้มากกว่าค่าเฉลี่ยของเพลงและค่าแนวโน้มของเพลงมากกว่าค่าเฉลี่ยของผู้ใช้จะใช้วิธีการทำนายค่าคะแนนดังสมการ (8)

$$p_{u_i,m_j} = \max(\bar{v}_{u_i} + ib_{m_j}, \bar{v}_{.m_j} + ub_{u_i}) \quad (8)$$



ภาพที่ 2: ความสัมพันธ์ระหว่างค่าเฉลี่ย (วงกลม) และแนวโน้มของผู้ใช้ (ลูกศร)

กรณีที่ 2 ภาพที่ 2(b) ทำนายจากแนวโน้มความชอบที่เป็นลบทั้งของผู้ใช้และเพลง ดังนั้นถ้าค่าแนวโน้มของผู้ใช้น้อยกว่าค่าเฉลี่ยของเพลงและค่าแนวโน้มของเพลงน้อยกว่าค่าเฉลี่ยของผู้ใช้จะใช้วิธีการทำนายค่าคะแนนดังสมการ(9)

$$p_{u_i,m_j} = \min(\bar{v}_{u_i} + ib_{m_j}, \bar{v}_{.m_j} + ub_{u_i}) \quad (9)$$

กรณีที่ 3 ภาพที่ 2(c) ทำนายจากแนวโน้มความชอบของเพลงที่เป็นบวกและผู้ใช้ที่มีแนวโน้มความชอบเป็นลบ โดยที่ถ้าค่าแนวโน้มของผู้ใช้น้อยกว่าค่าเฉลี่ยของเพลงและค่าแนวโน้มของเพลงมากกว่าค่าเฉลี่ยของผู้ใช้จะใช้วิธีการทำนายค่าคะแนนดังสมการ(10)

$$p_{u_i,m_j} = \min(\max(\bar{v}_{u_i}, (\bar{v}_{.m_j} + ub_{u_i}) \times (\alpha + (\bar{v}_{u_i} + ib_{m_j})(1-\alpha))), \bar{v}_{.m_j}) \quad (10)$$

โดยที่ α แสดงถึงค่าความเชื่อมั่นในการทำนายของค่าเฉลี่ยของผู้ใช้และเพลง

กรณีที่ 4 ภาพที่ 2(d) มีแนวโน้มความชอบที่แตกต่างกันอย่างชัดเจนซึ่งแนวโน้มความชอบของทั้งผู้ใช้และสินค้ามีแนวโน้มที่ตรงกันข้ามกัน โดยที่ถ้าค่าแนวโน้มของเพลงมากกว่าค่าเฉลี่ยของผู้ใช้และค่าแนวโน้มของผู้ใช้น้อยกว่าค่าเฉลี่ยของเพลงและค่าเฉลี่ยของผู้ใช้มากกว่าค่าเฉลี่ยของเพลงจะใช้วิธีการทำนายค่าคะแนนดังสมการ(11)

$$p_{u_i,m_j} = \bar{v}_{.m_j} \times \alpha + \bar{v}_{u_i} (1-\alpha) \quad (11)$$

3.5 การวัดประสิทธิภาพในการทำนาย

การวัดประสิทธิภาพของขั้นตอนวิธีการทำนายค่าคะแนนงานวิจัยนี้ใช้ตัววัดความแม่นยำทางสถิติที่เรียกว่า Mean Absolute Error (MAE)[12]ซึ่งเป็นการวัดค่าคะแนนความชอบที่ทำนายได้จากระบบมีความแตกต่างจากค่าคะแนนที่ผู้ใช้เคยให้ไว้เท่าไร โดยสูตรการคำนวณ MAE เป็นไปตามสมการที่ (12) และ (13)

$$MAE_{u_i} = \frac{\sum_{j=1}^n |v_{u_i,m_j} - pv_{u_i,m_j}|}{n} \quad (12)$$

$$MAE = \frac{\sum_{i=1}^N MAE_{u_i}}{N} \quad (13)$$

โดยที่ v_{u_i,m_j} คือคะแนนความชอบของเพลง m_j ที่ผู้ใช้ u_i ให้ไว้

pv_{u_i,m_j} คือคะแนนความชอบของเพลง m_j ที่ผู้ใช้ u_i ที่ได้จากการทำนาย

n คือ จำนวนเพลงที่ผู้ใช้ให้คะแนนไว้

N คือ จำนวนผู้ใช้

4. ผลการทดลอง

การวัดประสิทธิภาพในงานวิจัยนี้เป็นการเปรียบเทียบประสิทธิภาพระหว่างวิธีการที่นำเสนอ TBMS กับวิธีการกรองร่วมแบบดั้งเดิมที่พิจารณาผู้ใช้เป็นหลัก โดยใช้ตัววัดความคล้ายคลึงหลายตัววัดแบบต่างๆ โดยผลการทดลองเปรียบเทียบประสิทธิภาพมีรายละเอียดดังภาพที่ 3 ซึ่งพบว่าวิธีการที่นำเสนอ TBMS ให้ค่า MAE เฉลี่ยจากทุกชุด Micro Profiles คือ 0.5638 และวิธีการกรองร่วมที่พิจารณาผู้ใช้เป็นหลักในแต่ละความคล้ายคลึงต่างๆ และพิจารณาจำนวนเพื่อนบ้านเท่ากับ 10% ของจำนวนผู้ใช้ทั้งหมด พบว่าวิธีการกรองร่วมที่ใช้การวัดความคล้ายคลึงแบบโคไซน์ให้ค่า MAE คือ 0.6201 ใช้การวัดความคล้ายคลึงแบบสัมพันธ์สหสัมพันธ์ของเพียร์สันให้ค่า MAE คือ 0.6292 และใช้การวัดความคล้ายคลึงแบบโคไซน์ปรับแก้ให้ค่า MAE คือ 0.6166 ซึ่งสามารถสรุปได้ว่าวิธีการที่นำเสนอ TBMS มีประสิทธิภาพมากกว่าวิธีการกรองร่วมที่พิจารณาจากผู้ใช้เป็นหลัก

นอกจากการวัดความประสิทธิภาพของความแม่นยำในการทำนายแล้ว งานวิจัยนี้ยังวัดประสิทธิภาพในเชิงเวลาที่ใช้ในการ

ประมวลผลของแต่ละวิธีการอีกด้วย โดยทดลองบนเครื่องคอมพิวเตอร์ที่ใช้ซีพียูความเร็ว 3.20 GHz มีหน่วยความจำหลัก 8 GB ฮาร์ดดิสก์ 2 TB และภาษาโปรแกรมไพธอน (Python) ในการพัฒนาโปรแกรมดังแสดงผลในตารางที่ 5 ถึง 9

ภาพที่ 3: การเปรียบเทียบค่า MAE ระหว่างวิธีการที่นำเสนอ TBMS และวิธีการกรอกร่วมที่พิจารณาจากผู้ใช้เป็นหลัก

ตารางที่ 5: เวลาที่ใช้ในการประมวลผลสำหรับสร้าง Similarity Matrix

วิธีการ	เวลาที่ใช้ประมวลผล (วินาที)
Cosine Similarity (สมการที่ 1)	8326
Adjust Cosine Similarity (สมการที่ 2)	6027
Pearson correlation coefficient (สมการที่ 3)	6133
TBMS (ขั้นตอน 3.1+3.2+3.3)	171

ตารางที่ 6: เวลาที่ใช้ในการประมวลผลสำหรับขั้นตอนการทำนาย

วิธีการ	การทำนาย (วินาที)
Cosine Similarity (สมการที่ 4)	541
Adjust Cosine Similarity (สมการที่ 4)	587
Pearson correlation coefficient (สมการที่ 4)	583
TBMS (ขั้นตอน 3.4)	42

จากตารางที่ 5 วิธีการกรอกร่วมแบบดั้งเดิมที่ใช้วิธีการสร้าง Similarity Matrix แบบต่างๆ จะใช้เวลาในการประมวลผลสูงมากกว่าวิธีการ TBMS ที่นำเสนอในงานวิจัยนี้หลายเท่า สอดคล้องกับผลการทดลองในตารางที่ 6 ที่แสดงเวลาที่ใช้ในการทำนาย พบว่าวิธีการ TBMS มีเวลาในการประมวลผลที่เร็วกว่าทุกวิธีการเช่นกัน และตารางที่ 7 แสดงผลการทดลองของเวลาที่ใช้ในการทำนายของวิธีการ TBMS ซึ่งจากผลลัพธ์ที่

แสดงในตารางพบว่าเวลาในการทำนายจะลดลงเมื่อแบ่งกลุ่มของข้อมูลในการฟังเพลงของผู้ใช้ออกเป็น Micro Profiles เสียก่อน และผลจากแบ่งข้อมูลเป็น Micro Profiles นี้ทำให้มีค่า MAE เพียง 0.5638 และต่ำกว่าทั้งสามวิธีการที่นำมาเปรียบเทียบ (ภาพที่ 3)

ตารางที่ 7: เวลาในการทำนายของวิธีการ TBMS (ขั้นตอนที่ 3.4) โดยใช้ข้อมูลนำเข้าที่แบ่งเป็น Micro Profiles

วิธีการ TBMS		
ข้อมูลที่ใช้ในการทดลอง		เวลา (วินาที)
ไม่แบ่ง Micro Profiles	ข้อมูลทั้งหมด	42
แบ่ง Micro Profiles	ตอนเช้า	25
	ตอนกลางวัน	15
	ตอนเย็น	20
	วันหยุด	21
	วันธรรมดา	32

ตารางที่ 8: ประสิทธิภาพเชิงเวลาสำหรับสร้าง Similarity Matrix เปรียบเทียบกับวิธีการ TBMS โดยที่ m คือจำนวนผู้ใช้และ n คือจำนวนเพลง

วิธีการ	ประสิทธิภาพ
Cosine Similarity (สมการที่ 1)	$O(m^2 n)$
Adjust Cosine Similarity (สมการที่ 2)	$O(m^2 n)$
Pearson Similarity (สมการที่ 3)	$O(m^2 n)$
TBMS (ขั้นตอน 3.1+3.2+3.3)	$O(mn)$

ตารางที่ 9: ประสิทธิภาพเชิงเวลาสำหรับทำนายโดยที่ m คือจำนวนผู้ใช้และ n คือจำนวนเพลง

วิธีการ	การทำนาย
Cosine Similarity (สมการที่ 4)	$O(mn)$
Adjust Cosine Similarity (สมการที่ 4)	$O(mn)$
Pearson Similarity (สมการที่ 4)	$O(mn)$
TBMS (ขั้นตอน 3.4)	$O(1)$

จากตารางที่ 8 พบว่าประสิทธิภาพเชิงเวลาสำหรับสร้าง Similarity Matrix แบบต่างๆ ตามสมการที่ (1), (2) และ (3) เป็น $O(m^2 n)$ ส่วนวิธีการที่นำเสนอ TBMS มีประสิทธิภาพเชิงเวลาในขั้นตอนที่ 3.1 เป็น $O(mn)$ ขั้นตอนที่ 3.2 เป็น $O(mn)$ และ

ขั้นตอนที่ 3.3 เป็น $O(mn)$ ดังนั้นประสิทธิภาพโดยรวมทั้ง 3 ขั้นตอนจึงเป็น $O(mn)$ ซึ่งมีประสิทธิภาพดีกว่าทั้งสามวิธีการที่ใช้เปรียบเทียบ และจากตารางที่ 9 แสดงประสิทธิภาพเชิงเวลาของการทำนายพบว่า วิธีการที่เสนอ TBMS ซึ่งประยุกต์ใช้หลักการของวิธีการ “Tendencies Based Algorithm” ในการทำนาย โดยในงานวิจัย [8] ได้ทำการทดลอง และหาค่าประสิทธิภาพในการทำนายของวิธีการ “Tendencies Based Algorithm” เป็น $O(1)$ ดังนั้นวิธีการ TBMS ที่นำเสนอนี้ จึงมีประสิทธิภาพเชิงเวลาของการทำนายเป็น $O(1)$ เช่นเดียวกัน ซึ่งเป็นประสิทธิภาพที่ดีกว่าทั้งสามวิธีการที่เปรียบเทียบอย่างมาก

5. สรุปผลการทดลอง

งานวิจัยนี้ นำเสนอขั้นตอนวิธีสำหรับระบบแนะนำข้อมูลเพลงชื่อว่า “Tendencies Based Algorithm with Micro Profiles & Sparsity Problem (TBMS)” ซึ่งประยุกต์ใช้วิธีการของ Tendencies Based Algorithm มาใช้ และปรับปรุงขั้นตอนการทำนายค่าคะแนนความชอบในระบบแนะนำเพลงให้มีความแม่นยำในการทำนายมากขึ้น โดยการใช้ Micro Profiles อีกทั้งยังแก้ปัญหาที่เป็นค่าคะแนนที่ว่าง โดยการเติมค่าว่างก่อนทำนาย ซึ่งจากผลการทดลองแสดงให้เห็นว่าวิธีการ TBMS นี้ให้ผลการทำนายความชอบที่แม่นยำกว่าวิธีการกรองร่วมแบบดั้งเดิม นอกจากนี้ประสิทธิภาพในเชิงเวลาของวิธีการ TBMS ก็ยังมีความรวดเร็วมากกว่าวิธีการกรองร่วมแบบดั้งเดิมอีกด้วย

แต่อย่างไรก็ตามงานวิจัยนี้ยังมีแนวทางในการปรับปรุงให้วิธีการทำนายค่าคะแนนให้มีความแม่นยำมากยิ่งขึ้น โดยการปรับปรุงขั้นตอนวิธีในการเติมค่าคะแนนที่ว่างที่พิจารณาจากปัจจัยอื่นที่เกี่ยวข้อง นอกเหนือจากพิจารณาเพียงแค่แนวเพลงเพียงอย่างเดียว รวมทั้งการแบ่งช่วงเวลา Micro Profiles แบบช่วงเวลาของผู้ใช้เฉพาะบุคคลมากขึ้น

เอกสารอ้างอิง

- [1] K.Marius and F. Ricci, "Contextual music information retrieval and recommendation:state of the art and challenges," *Computer Science Review*, Vol.6, pp. 89–119, 2012.
- [2] J.Bobadilla, F.Ortega, A.Hernando, A.Gutierrez, "Recommender systems survey" ,*Knowledge-Based Systems*, Vol.46, pp.109-132, July, 2013.
- [3] C.Laurent, J.Kris, F.Francoise, M.Frank, *State-of-the-Art Recommender Systems* , IGI Global, 2009.
- [4] C. H. Park and M. Kahng, "Temporal Dynamics in Music Listening Behavior:A Case Study of Online Music Service ",*The 9th International Conference on Computer and Information Science (ACIS 2010)*, pp.573-578, 18-20 August,2010.
- [5] B. Linas and A. Xavier, "Towards time-dependant recommendation based on implicit feedback," *CARS-2009*, New York, USA, October 25, 2009.
- [6] S. Antonietta, A. Albert, D. Helenca, E. Gregorio, and V. Paulo, "My Personal Media Entertainer: Context-Adaptive Content Recommendation and Delivery ", *ACM Workshop on Social, Adaptive and Personalized Multimedia Interaction and Access*, pp.33-36, 2010.
- [7] T. Cebrían, M. Planagumà, P. Villegas, and X. Amatriain, "Music Recommendations with Temporal Context Awareness," *Recommender Systems (RecSys2010)*, pp.349-352, 2010.
- [8] C. Fidel, C. Victor, F. Diego and F. Vreixo, "Comparison of Collaborative Filtering Algorithms." *ACM Transaction on the Web*, Vol. 5, No.1, Febuary, 2011.
- [9] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Item-Based Collaborative Filtering Recommendation Algorithms," *Proceedings of the 10th international conference on World Wide Web(www2001)*, Hong Kong, May 1-5, 2001.
- [10] C.Oscar. Last fm Dataset. [online]. <http://www.dtic.upf.edu/~ocelma/MusicRecommendationDataset/lastfm-1K.html> (วันที่ค้นข้อมูล : 1 กันยายน 2556)
- [11] B. Xu, J. Bu, C. Chen, and D. Cai, "An Exploration of Improving Collaborative Recommender Systems via User-Item Subgroups " *WWW2012*, Lyon, France, Apr 16-20, 2012.
- [12] C. J. Willmott and K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," *CLIMATE RESEARCH*, Vol. 30, pp.79-82, Dec 19, 2005.

การพยากรณ์ดัชนีราคาผู้บริโภคด้วยเทคนิคการทำเหมืองข้อมูลกรณีศึกษาจังหวัดสระแก้ว Consumer Prices Index Prediction using Data Mining Techniques: A Case Study of Sakaeo Province

กิตติศักดิ์ อ่อนอ่อน (K. Onuean)¹ และพยุ่ง มีสีจ (P. Meesad)²

¹สาขาวิชาคอมพิวเตอร์ธุรกิจ คณะวิทยาศาสตร์และสังคมศาสตร์ มหาวิทยาลัยบูรพา วิทยาเขตสระแก้ว

²คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

kittisak@buu.ac.th and pym@kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อสร้างรูปแบบการพยากรณ์ดัชนีราคาผู้บริโภคของจังหวัดสระแก้ว ด้วยเทคนิคเหมืองข้อมูล ซึ่งผู้วิจัยได้นำข้อมูลดัชนีราคาผู้บริโภคของจังหวัดสระแก้ว ช่วงปี พ.ศ. 2545-2555 มาวิเคราะห์ตามหมวดหมู่ของดัชนีราคาผู้บริโภคจำนวน 8 หมวดหมู่ จำนวน 80 แอททริบิวต์ โดยนำ 4 อัลกอริทึม ในการทดสอบข้อมูล ได้แก่ Classification tree (CART), Classification tree (C4.5), Rule Base (CN2) และ Naïve Bayes โดยใช้ค่าความถูกต้องในการจำแนกกลุ่มข้อมูล สูงสุดในการนำอัลกอริทึมมาใช้ ซึ่งผลที่ได้จากการวิจัยได้ อัลกอริทึม Rule Base (CN2) ให้ค่าถูกต้องสูงสุดเท่ากับ 0.97 และมีกฎทั้งหมด 16 กฎ และนำกฎที่ได้มาพยากรณ์ดัชนีราคาผู้บริโภคในเดือนถัดไป เพื่อให้ผู้บริหารใช้ข้อมูลในการสนับสนุนการตัดสินใจในการสร้างกลยุทธ์ในการบริหารจัดการ และใช้เป็นข้อมูลในการประกอบการตัดสินใจในการกำหนดราคาค่าแรงของประชาชนในจังหวัดสระแก้ว

คำสำคัญ: เหมืองข้อมูล ดัชนีราคาผู้บริโภค การจำแนกข้อมูล การพยากรณ์

Abstract

The objective of this research is to build prediction models for Consumer Price Index (CPI) of Sakaeo province by using data mining techniques. The 2002-2012 CPI's data obtained from Sakaeo province was used, which can be analyzed into 8 categories 80 attributes and the data mining models used were 4 different algorithms: Classification And Regression Tree (CART), : Classification Tree (C4.5), Rule Base (CN2),

and Naïve Bayes. The research results show that the Rule Base (CN2) algorithm resulted in 16 rules outperforms the other remaining algorithms; the accuracy was 0.97. The obtained rules were used to predict CPI of the next month in which it can be applied in decision support as a management strategy using data to assign labor prices of people in Sakaeo province.

Keyword: Data mining, Consumer Price Index, Data Classification, Prediction

1. บทนำ

จังหวัดสระแก้ว เป็นจังหวัดที่มีอาณาเขตติดต่อกับประเทศกัมพูชา โดยถือว่าเป็นจังหวัดที่มีความสำคัญจังหวัดหนึ่ง เนื่องจากเป็นจังหวัดที่เป็นประตูสู่อินโดจีน ประชาชนในจังหวัดมีการประกอบอาชีพต่างๆ เช่น ธุรกิจค้าขาย ธุรกิจอุตสาหกรรม การเกษตร การขนส่งระหว่างประเทศ เป็นต้น ดังนั้นจากสภาพการณ์ดังกล่าว จึงมีแรงงานต่างประเทศ เข้ามาใช้แรงงานในจังหวัด ทำให้ราคาค่าแรงมีการเปลี่ยนแปลง และเกิดการแข่งขันในตลาดแรงงานของประชาชนในจังหวัดสระแก้วซึ่งจากงานวิจัย [1] การวิเคราะห์หาปัจจัยที่ส่งผลกระทบต่อดัชนีราคาผู้บริโภคเพื่อรองรับเขตเศรษฐกิจพิเศษด้วยเทคนิคการทำเหมืองข้อมูล เป็นงานวิจัยที่วิเคราะห์หาปัจจัยหลักที่ส่งผลกระทบต่อดัชนีราคาผู้บริโภค (Consumer Price Index :CPI) [2] ที่เป็นเครื่องมือทางสถิติที่ใช้ในการแสดงถึงภาวะเศรษฐกิจของประชาชนในจังหวัดต่างๆ และรวบรวมจังหวัดต่างๆ ออกมาเป็นดัชนีราคาผู้บริโภคของประเทศ โดยมีการจัดเก็บข้อมูลออกเป็น 20 ชุดข้อมูล และแบ่งออกเป็น 8 หมวดหมู่ ได้แก่ หมวดอาหารและเครื่องดื่ม หมวดเครื่องนุ่งห่ม

และรองเท้า หมวดเกษตรสถาน หมวดการตรวจรักษาและบริการส่วนบุคคล หมวดพาหนะ การขนส่ง และการสื่อสาร หมวดการบันเทิงการอ่านและการศึกษา หมวดยาสูบและเครื่องดื่มมีแอลกอฮอล์ และดัชนีราคาผู้บริโภคพื้นฐาน ซึ่งจากงานวิจัยยังไม่สามารถพยากรณ์ดัชนีราคาผู้บริโภคในอนาคตได้

ดังนั้นในงานวิจัยนี้ได้นำข้อมูลดัชนีราคาผู้บริโภคของจังหวัดสระแก้ว เป็นกรณีศึกษา เพื่อนำดัชนีราคาผู้บริโภคมาใช้ในการสร้างโมเดลพยากรณ์ดัชนีราคาผู้บริโภคในเดือนถัดไป โดยมีการแปลงผลข้อมูลแบ่งออกเป็น 3 กลุ่ม ได้แก่ 1=อัตราเจริญเติบโตของดัชนีราคาผู้บริโภคที่ลดลง 2=อัตราเจริญเติบโตของดัชนีราคาผู้บริโภคที่คงที่ และ 3=อัตราเจริญเติบโตของดัชนีราคาผู้บริโภคที่เพิ่มขึ้น จากนั้นนำข้อมูลที่นำมาใช้อัลกอริทึมจำนวน 4 อัลกอริทึม มาทำการสร้างโมเดล ได้แก่ Decision tree (Information Gain), Decision tree (GINI Index), Rule base และ Naïve Bayes มาใช้ในการวิจัย และนำกฎที่ได้จากการสร้างโมเดลที่ให้ค่าความถูกต้องสูงสุดมาประยุกต์ใช้ในการสร้างระบบพยากรณ์ดัชนีราคาผู้บริโภค เพื่อเป็นการได้สารสนเทศจากการพยากรณ์มาช่วยสนับสนุนในการตัดสินใจให้กับผู้บริหารของจังหวัดในการประมาณการอัตราค่าแรงของประชาชนในจังหวัดสระแก้ว

2. งานวิจัยที่เกี่ยวข้อง

2.1 ข้อมูลพื้นฐานจังหวัดสระแก้ว

จังหวัดสระแก้ว [3] ได้ยกฐานะขึ้นเป็นกิ่งอำเภอชื่อว่า “กิ่งอำเภอสระแก้ว” ในปี พ.ศ.2452 ขึ้นอยู่ในการปกครองของอำเภอกบินทร์ โดยใช้ชื่อสระน้ำ เป็นชื่อกิ่งอำเภอชื่อว่า “อำเภอสระแก้ว” ขึ้นอยู่ในการปกครองของจังหวัดปราจีนบุรี และต่อมาเมื่อวันที่ 1 ธันวาคม 2536 ได้มีพระราชบัญญัติจัดตั้งจังหวัดสระแก้วขึ้น ประกาศในพระราชกิจจานุเบกษา ฉบับพิเศษ เล่มที่ 110 ตอนที่ 125 ลงวันที่ 2 กันยายน 2536 เป็นผลให้จังหวัดสระแก้วได้เปิดทำการในวันที่ 1 ธันวาคม 2536 โดยเป็นจังหวัดที่ 74 ของประเทศไทย โดยจังหวัดสระแก้วได้มีการกำหนดวิสัยทัศน์ของจังหวัด ได้แก่ [4] “ศูนย์กลางโลจิสติกส์ และแหล่งท่องเที่ยวเชิงนิเวศของอินโดจีน ถิ่นพืชพลังงานอาหารปลอดภัย” เป็นยุทธศาสตร์การพัฒนาจังหวัดสระแก้ว ที่เป็นประเด็นมาสู่ประเด็นยุทธศาสตร์ของจังหวัดสระแก้วทั้ง 5

ประเด็นยุทธศาสตร์ โดยแต่ละประเด็นยุทธศาสตร์ของจังหวัดสระแก้วนั้นผู้บริหารจังหวัดได้สังเกตเห็นถึงความสำคัญของดัชนีราคาผู้บริโภคในจังหวัดที่ส่งผลต่อสถานะเศรษฐกิจภายในจังหวัดสระแก้ว

2.2 ดัชนีราคาผู้บริโภค

ดัชนีราคาผู้บริโภค หรือ Consumer Price Index (CPI) [2] เป็นเครื่องมือทางสถิติที่ใช้วัดการเปลี่ยนแปลงราคาขายปลีกของสินค้าและบริการในปริมาณที่เท่ากันในช่วงระยะเวลาหนึ่งเทียบกับราคาสินค้าอย่างเดียวกันในช่วงเวลาตั้งต้น หรือเรียกว่าปีฐาน (Base Year) ซึ่งเมื่อดัชนีราคาผู้บริโภคเพิ่มขึ้นจากเดิมแสดงว่าสถานะเศรษฐกิจนั้นมีการเกิดเงินเฟ้อขึ้น กล่าวได้คือมูลค่าของสินค้ามีมูลค่าเพิ่มขึ้นจากเดิม โดยดัชนีราคาผู้บริโภคสามารถนำมาใช้ในการวัดค่าครองชีพของประชาชนทุกระดับรายได้ วัดภาวะเงินเฟ้อระดับจังหวัดและประเทศ เป็นมาตรฐานในการปรับค่าจ้าง เงินเดือน บำเหน็จ บำนาญ และเงินช่วยเหลือสวัสดิการสังคม เป็นแนวทางการปรับค่าจ้างแรงงาน เป็นราคาสัญญาก่อสร้าง และราคาที่จะมีการซื้อขายในระยะยาว ใช้คำนวณค่าของเงินที่แท้จริง (Real Value) ของประชาชน ใช้เป็นข้อมูลประกอบการกำหนดนโยบายการเงินและการคลังของประเทศ ใช้เป็นแนวทางในการวิจัยพยากรณ์การตลาดและราคาสินค้าต่างๆ ใช้วัดผลความสำเร็จในการพัฒนาชนบทซึ่งมีข้อมูลทั้งหมด 20 ประเภท ซึ่งถูกจัดกลุ่มออกเป็น 8 หมวดหมู่ ได้แก่ หมวดอาหารและเครื่องดื่ม หมวดเครื่องนุ่งห่มและรองเท้า หมวดเกษตรสถาน หมวดการตรวจรักษาและบริการส่วนบุคคล หมวดพาหนะ การขนส่ง และการสื่อสาร หมวดการบันเทิงการอ่านและการศึกษา หมวดยาสูบและเครื่องดื่มมีแอลกอฮอล์ และดัชนีราคาผู้บริโภคพื้นฐาน โดยงานวิจัยได้นำดัชนีราคา [5] ผู้บริโภคจังหวัดสระแก้วตั้งแต่ปี พ.ศ. 2545 ถึง พ.ศ. 2555 มาใช้ในการสร้างโมเดล ซึ่งได้มีการนำข้อมูลดัชนีราคาผู้บริโภคมาใช้ในการจัดทำงานวิจัยในการหาปัจจัยที่ส่งผลต่อดัชนีราคาผู้บริโภคของจังหวัดสระแก้ว ด้วยเทคนิคการทำเหมืองข้อมูล [1] โดยการนำกฎที่ได้จากโมเดลมาคำนวณหาค่าความถี่ของหมวดหมู่ของดัชนีราคาผู้บริโภคที่มีผลต่อดัชนีราคาผู้บริกรวมของจังหวัดสระแก้ว

2.3 เทคนิคการทำเหมืองข้อมูล

การทำเหมืองข้อมูล (Data Mining) เป็นเทคนิควิธีการเพื่อค้นหารูปแบบของข้อมูลในปริมาณมากแบบอัตโนมัติ เพื่อให้ได้ซึ่ง รูปแบบ แนวทาง และความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูล โดยอาศัยการพยากรณ์จากสถิติ การเรียนรู้ และการรู้จำ [6] ซึ่งในงานวิจัยได้นำ 4 อัลกอริทึมที่นิยมใช้ในการงานวิจัยเกี่ยวกับการทำเหมืองข้อมูลมาใช้ดังนี้

ต้นไม้ตัดสินใจ (Decision Tree: DT) เป็นแบบจำลองที่นิยมใช้ในกระบวนการทำเหมืองข้อมูล (Data Mining) และง่ายในการจำแนกข้อมูลออกเป็นคลาสต่างๆ ซึ่งมีลักษณะโครงสร้างการตัดสินใจในลักษณะโครงสร้างขององค์กร (Organization chart) ที่มีลักษณะแบบบนลงล่าง โดยมี root node อยู่บนสุดใช้ในการจำแนกข้อมูลเริ่มแรกเมื่อทำการจำแนกข้อมูล ซึ่งการจัดทำต้นไม้ตัดสินใจนี้จะทำงานแบบเรียนรู้แบบมีครู (Supervised Learning) โดยในการสร้างโมเดลจะทำการแบ่งข้อมูลออกเป็น 2 ชุด คือ ชุดในการเรียนรู้และสร้างโมเดล (Training Data) และชุดข้อมูลในการทดสอบความถูกต้องของโมเดล (Test Model) โดยจะนำโมเดลที่ได้จากการสร้างขึ้นมาทำการทดสอบข้อมูลที่ต้องการพยากรณ์ ซึ่งประกอบด้วย อัลกอริทึมต่างๆ เช่น Hunt's Algorithm, CART, ID3, C4.5, SLIQ และ SPRINT เป็นต้นซึ่งในงานวิจัยนี้จะนำอัลกอริทึมแบบ CART และ C4.5 มาใช้ในการสร้างโมเดลโดยใช้ GINI index เป็นกระบวนการจำแนกแบบ CART [7] ที่เป็นกระบวนการที่เหมาะสมสำหรับการจำแนกข้อมูลที่เป็นแบบสองทาง โดยมีการจำแนกข้อมูลแบบสมการถดถอยและการจำแนกข้อมูล ซึ่งจะมีการแยกส่วนย่อยหรือ โหนดลูกออก ซึ่งจะใช้การหาค่า GINI ในการแบ่งข้อมูล และ Information Gain เป็นกระบวนการจำแนกแบบ C4.5 Algorithm เป็นอัลกอริทึมที่ได้จากกลไกในการเรียนรู้ [8] ใช้ในการเลือก Attribute ที่ดีที่สุดมาสร้างเป็น node ในการสร้าง Decision Tree โดยอาจทำการหาได้จากค่า information gain ในการแยก node ที่ดีที่สุด โดยจะเริ่มทำการหาตั้งแต่การหาค่า Entropy โดยสามารถนำค่า Entropy ที่ได้มาทำการคำนวณหาค่า Information Gain

กฎพื้นฐาน (Rule Base: CN2) [9] เป็นกฎในจำแนกข้อมูลที่ดี ในการทำงาน CN2 จะทำการหากฎใหม่เพิ่มขึ้น จากชุดข้อมูลที่นำมาทดสอบ โดยที่ CN2 เป็นลักษณะของการหา

แบบทางตรงโดยจะเริ่มค้นจากยังไม่มีความรู้ใดๆ และจะเรียนรู้กฎเพื่อนำตัวอย่างข้อมูลมาทำการทดสอบความถูกต้องของข้อมูลไปเรื่อยๆ ซึ่งสามารถสรุปได้คือ 1) เริ่มต้นจากยังไม่มีความรู้ใดๆ ที่เกิดขึ้น 2) เรียนรู้กฎที่เกิดขึ้นจากชุดข้อมูลทดสอบ (Training Data) 3) เมื่อได้กฎทำการกำจัดข้อมูลในข้อมูลทดสอบที่ครอบคลุมกฎที่ได้ และ 4) จัดทำขั้นตอน 2 และ 3 จนกระทั่งทั้งหมด

นาอิวเบย์ (Naïve Bayes) เป็นอัลกอริทึมที่ในการจำแนกข้อมูลที่กำหนดเป็นแต่ละหมวดหมู่โดยอาศัยหลักการความน่าจะเป็นของกลุ่มต่างๆ [10] ที่สามารถเกิดขึ้นได้จากแอตทริบิวต์ทั้งหมดว่าขึ้นตรงกับเงื่อนไขของกลุ่มข้อมูลในแต่ละกลุ่ม ซึ่งถ้าแอตทริบิวต์ที่ทดสอบนั้นตรงกับเงื่อนไขของกลุ่มนั้น จะนำมาเป็นสมาชิกของกลุ่มนั้น โดยในการจำแนกการจัดทำต้องดำเนินการหาค่าความน่าจะเป็นที่จะเกิดขึ้นในแต่ละ Class ว่าเป็นเท่าไร โดยในกระบวนการหาว่าจะอยู่ Class ใดนั้นต้องดำเนินการเริ่มต้นตั้งแต่การหาค่าความน่าจะเป็นของ Class แต่ละ Class ที่เกิดขึ้นในการจำแนกกลุ่ม จากนั้น หาค่าความน่าจะเป็นของทุกๆ กรณีที่คาดว่าจะเกิด Class นั้นๆ และทำการนำมาคูณกัน ได้ค่าความน่าจะเป็นที่จะเกิดขึ้นในแต่ละ Class ซึ่งถ้าค่าความน่าจะเป็นใน Class ใดๆ มีค่ามากที่สุด ถือว่าข้อมูล Record นั้นเป็น Class ดังกล่าว

3. ขั้นตอนวิธีการดำเนินการวิจัย

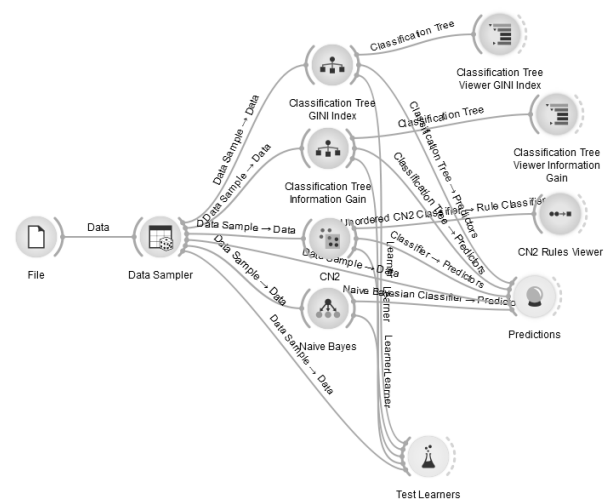
การดำเนินงานวิจัยเริ่มต้นจากการนำข้อมูลดัชนีราคาผู้บริโภคจังหวัดสระแก้วตั้งแต่ปี พ.ศ. 2545 ถึงปี พ.ศ. 2555 รวมจำนวน 11 ปีมาใช้ในการทดลองงานวิจัย ซึ่งได้นำค่าดัชนีราคาผู้บริโภคจำนวน 8 หมวดหมู่ [2] โดยกำหนดให้ G1=หมวดอาหารและเครื่องดื่ม G2=หมวดเครื่องนุ่งห่มและรองเท้า G3=หมวดเคหสถาน G4=หมวดการตรวจรักษาและบริการส่วนบุคคล G5=หมวดพาหนะ การขนส่ง และการสื่อสาร G6=หมวดการบันเทิงการอ่านและการศึกษา G7=หมวดยาสูบและเครื่องดื่มมีแอลกอฮอล์ และ G8=ดัชนีราคาผู้บริโภคพื้นฐาน จากนั้นผู้จัดทำงานวิจัยได้นำมาจัดทำข้อมูลให้อยู่ในรูปแบบของงานวิจัยโดยกำหนดให้ ดัชนีราคาผู้บริโภคที่เพิ่มขึ้น=3 ดัชนีราคาผู้บริโภคคงที่=2 และดัชนีราคาผู้บริโภคลดลง=1 โดยใช้ปี พ.ศ. 2545 เป็นปีฐานในการคำนวณ ซึ่งประกอบด้วย 80

attribute และ 1 target class ซึ่ง target class ที่งานวิจัยได้นำมาใช้ คือ ค่าของดัชนีราคาผู้บริโภคในเดือนถัดไป เทียบกับดัชนีราคาผู้บริโภคในเดือนปัจจุบัน ซึ่งแบ่งออกเป็น 3 ประเภท ได้แก่ ดัชนีราคาผู้บริโภคเดือนถัดไปลดลง=1 ดัชนีราคาผู้บริโภคเดือนถัดไปคงที่=2 และ ดัชนีราคาผู้บริโภคเดือนถัดไปเพิ่มขึ้น=3 โดยมีนำค่าดัชนีราคาผู้บริโภคมาจัดรูปแบบดังตารางที่ 1

ตารางที่ 1: ตารางอธิบายตัวอย่างรายละเอียดข้อมูลนำเข้าโมเดล

ชื่อ	รายละเอียด
GX0	เดือนปัจจุบันลบด้วยเดือนก่อนหน้า
GX1	เดือนปัจจุบันลบด้วยสามเดือนย้อนหลัง
GX2	เดือนปัจจุบันลบด้วยค่าเฉลี่ยสามเดือนย้อนหลัง
GX3	เดือนปัจจุบันลบด้วยหกเดือนย้อนหลัง
GX4	เดือนปัจจุบันลบด้วยค่าเฉลี่ยหกเดือนย้อนหลัง
GX5	เดือนปัจจุบันลบด้วยเก้าเดือนย้อนหลัง
GX6	เดือนปัจจุบันลบด้วยค่าเฉลี่ยเก้าเดือนย้อนหลัง
GX7	เดือนปัจจุบันลบด้วยสิบสองเดือนย้อนหลัง
GX8	เดือนปัจจุบันลบด้วยค่าเฉลี่ยสิบสองเดือนย้อนหลัง
GX9	เดือนปัจจุบันลบด้วยเดือนเดียวกันเมื่อปีที่แล้ว

จากตารางที่ 1 แสดงตัวอย่างรายละเอียดของข้อมูลโดยกำหนดให้ GX แทนดัชนีราคาผู้บริโภคตามหมวดต่างๆ รวม 8 หมวดหมู่ ๆ ละ 10 ตัวแปร เช่น G11 คือดัชนีราคาผู้บริโภคของหมวดอาหารและเครื่องดื่ม ณ เดือนปัจจุบันลบด้วยเดือนก่อนหน้า หากมีค่าเป็นบวก (เพิ่ม) ให้กำหนดค่าเท่ากับ 3 หากมีค่าเป็นศูนย์ (คงที่) ให้กำหนดค่าเท่ากับ 2 และหากมีค่าน้อยกว่าศูนย์ (ลด) ให้กำหนดค่าเท่ากับ 1 เป็นต้น โดยมีการคำนวณเริ่มต้นจากปี พ.ศ. 2546 เป็นต้นไป เนื่องจากปี พ.ศ. 2555 ต้องนำมาคำนวณเป็นปีฐาน จากนั้นเมื่อได้ข้อมูลจำนวน 81 แอททริบิวต์ (รวม Target Class) ทางผู้จัดทำงานวิจัยได้นำโปรแกรม Orange Canvas [11] ซึ่งเป็นเครื่องมือที่เข้ามาช่วยในการจัดทำเหมืองข้อมูล ซึ่งมีคุณสมบัติในการออกแบบข้อมูลและแสดงผลแผนผังการออกแบบที่ทำให้เข้าใจง่าย ซึ่งเริ่มต้นจากการนำข้อมูลที่ผ่านประมวลผลมาผ่านการทดสอบของอัลกอริทึม 4 อัลกอริทึม แสดงในโมเดลดังภาพที่ 1



ภาพที่ 1: ภาพโมเดลค่าทำนายโดยใช้โปรแกรม Orange Canvas

จากภาพที่ 1 ผู้วิจัยได้นำข้อมูลเข้าโดยเริ่มต้นจากการจัดการข้อมูลให้มีจำนวน 8 หมวดหมู่ จำนวน 80 attribute ที่ต้องการและกำหนดคลาสที่ต้องการ จากนั้น จึงทำการแบ่งข้อมูลออกเป็น 2 ชุด เพื่อใช้ในการสร้างโมเดล และใช้ในการทดสอบโมเดล โดยแบ่งชุดข้อมูลออกเป็น 80% สำหรับการสร้างโมเดล (training data) และ 20% สำหรับการทดสอบโมเดล (test data) ส่วนผลที่ได้จะนำมาเปรียบเทียบเพื่อหาประสิทธิภาพในการจำแนกในแง่ของความถูกต้องหรือ Classification Accuracy (CA) ซึ่งได้ค่าความแม่นยำในการจำแนก ดังแสดงในตารางที่ 2 ด้านล่าง

ตารางที่ 2: ตารางค่าความแม่นยำของแต่ละอัลกอริทึม

อัลกอริทึม	ความแม่นยำ
Classification Tree Gini index	0.93
Classification Tree Information Gain	0.93
<u>Rule bases (CN2)</u>	<u>0.97</u>
Naive Bayes	0.80

จากตารางที่ 2 แสดงค่าความแม่นยำของแต่ละอัลกอริทึมซึ่งแสดงให้เห็นว่า อัลกอริทึมที่ให้ความแม่นยำสูงสุดได้แก่ Rule bases (CN2) เท่ากับ 0.97 ซึ่งเป็นอัลกอริทึมที่มีวิธีการจากการเรียนรู้จากชุดข้อมูล อันดับต่อมาเป็น Classification Tree Gini index และ Classification Tree Information Gain เท่ากับ 0.93 และอันดับสุดท้าย Naive Bayes เท่ากับ 0.80 ดังนั้นจึงนำ

กฎที่ได้ของอัลกอริทึม Rule bases (CN2) มาใช้ในการพยากรณ์ การเพิ่ม คงที่ และลดลง ของดัชนีราคาผู้บริโภคต่อไป

4. ผลการดำเนินงานวิจัย

จากการดำเนินงานวิจัย ผู้จัดทำงานวิจัยสามารถได้กฎที่ได้ จากอัลกอริทึมจำนวน 16 กฎ ดังต่อไปนี้

IF G40=3 AND G33=1 AND G65=3 THEN GC=1

IF G63=1 AND G61=2 THEN GC=1

IF G79=1 AND G67=2 THEN GC=1

IF G82=1 AND G50=2 AND G33=3 THEN GC=1

IF G23=2 AND G37=3 AND G34=3 THEN GC=1

IF G31=2 AND G39=1 AND G20=2 THEN GC=1

IF G14=1 AND G42=1 AND G45=1 THEN GC=2

IF G64=2 AND G23=3 THEN GC=3

IF G82=2 AND G46=3 THEN GC=3

IF G80=1 AND G88=1 AND G84=1 THEN GC=3

IF G66=2 AND G50=3 THEN GC=3

IF G61=1 AND G30=2 AND G12=3 THEN GC=3

IF G41=2 AND G37=1 THEN GC=3

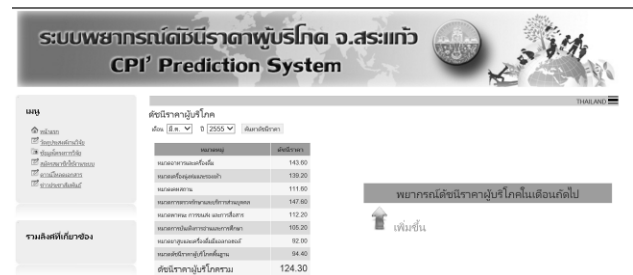
IF G32=1 AND G50=2 THEN GC=3

IF G61=1 AND G40=2 AND G81=3 THEN GC=3

IF G33=2 AND G11=1 THEN GC=3

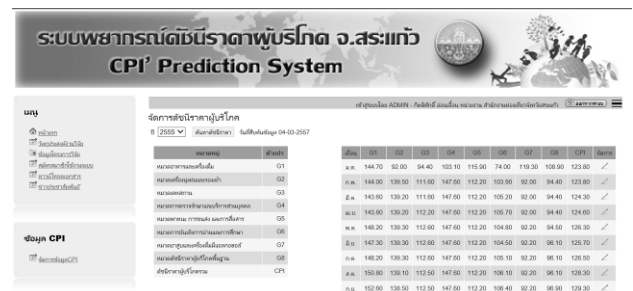
จากผลที่ได้ในการศึกษานี้ ทำให้ผู้บริหารสามารถพยากรณ์ ดัชนีราคาผู้บริโภคในเดือนถัดไปได้จากกฎที่ได้จากการทดลอง ได้ เช่น จากกฎ IF G40=3 AND G33=1 AND G65=3 THEN GC=1 คือ เมื่อดัชนีราคาผู้บริโภคของหมวดการตรวจรักษาและบริการส่วนบุคคล(G4)ในเดือนปัจจุบันเทียบกับเดือนที่แล้ว เพิ่มขึ้น และ ดัชนีราคาผู้บริโภคหมวดเคหสถาน(G3)ในเดือน ปัจจุบันเทียบกับดัชนีราคาผู้บริโภคหมวดเดือนย้อนหลังลดลง และดัชนีราคาหมวดการบันเทิงการอ่านและการศึกษา (G6)ในเดือนปัจจุบันเทียบกับดัชนีราคาผู้บริโภคเก่าเดือนย้อนหลังมีค่า เพิ่มขึ้น แสดงว่าดัชนีราคาผู้บริโภครวมในเดือนถัดไปมีโอกา สลดลง โดยเมื่อได้กฎต่างๆ ที่ได้จากการวิจัยทางผู้จัดทำงานวิจัย ได้พัฒนาระบบพยากรณ์ดัชนีราคาผู้บริโภค จังหวัดสระแก้ว ใน รูปแบบเว็บแอปพลิเคชัน ซึ่งประกอบด้วยการบันทึกข้อมูล CPI

ตามปีต่าง และการพยากรณ์ CPI ในเดือนถัดไปได้จากกฎ แสดงดังภาพด้านล่าง



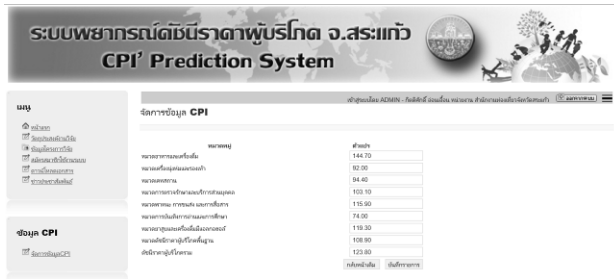
ภาพที่ 2: ภาพหน้าจอการพยากรณ์ CPI จังหวัดสระแก้ว

จากภาพที่ 2 ระบบพยากรณ์ดัชนีราคาผู้บริโภคจังหวัด สระแก้ว มีการแสดงผลในส่วนของการพยากรณ์ดัชนีราคา ผู้บริโภค ณ เดือนปัจจุบัน หรือ เดือนและปีที่ต้องการค้นหาการ พยากรณ์ได้จากระบบซึ่ง มีส่วนในการแสดงผล 2 ส่วนด้วยกัน คือ ค่าดัชนีราคาผู้บริโภคในเดือนนั้น และการพยากรณ์ดัชนี ราคาผู้บริโภครวมในเดือนถัดไป ซึ่งได้จากการบันทึกข้อมูล ดัชนีราคาผู้บริโภคตามเดือนต่างๆ แสดงดังภาพด้านล่าง



ภาพที่ 3: ภาพหน้าจอจัดการข้อมูลดัชนีราคาผู้บริโภค

จากภาพที่ 3 ภาพหน้าจอการจัดการข้อมูลดัชนีราคา ผู้บริโภค เป็นหน้าจอในการจัดการข้อมูลของดัชนีราคา ผู้บริโภคตามเดือนต่างๆ ซึ่งเมื่อผู้ใช้ต้องการแก้ไข หรือบันทึก ข้อมูลดัชนีราคาผู้บริโภค สามารถทำการเลือกที่สัญลักษณ์ เพิ่ม หรือแก้ไข ข้อมูล ระบบจะแสดงหน้าจอการเพิ่ม หรือแก้ไข ข้อมูลดังภาพที่ 4



ภาพที่ 4: ภาพหน้าจอการเพิ่มข้อมูลดัชนีราคาผู้บริโภค

จากภาพที่ 4 เมื่อผู้ใช้งานต้องการบันทึกข้อมูลสามารถทำการกรอกข้อมูลดัชนีราคาผู้บริโภคตามหมวดหมู่ต่างๆ ผ่านระบบ จากนั้นทำการเลือกปุ่มบันทึกการระบบจะทำการบันทึกข้อมูลในฐานข้อมูล ซึ่งจากความสามารถทำงานของระบบพยากรณ์ดัชนีราคาผู้บริโภคทำให้สามารถนำผลที่ได้ไปทำการกำหนดกลยุทธ์ในการดำเนินการเกี่ยวกับดัชนีราคาผู้บริโภค และสามารถนำข้อมูลจากการพยากรณ์ไปใช้ในการประกอบการตัดสินใจในการกำหนดราคาค่าแรง และค่าครองชีพของประชาชนจังหวัดสระแก้ว

5. สรุป

งานวิจัยนี้นำเสนอเทคนิคการพยากรณ์ดัชนีราคาผู้บริโภคซึ่งได้นำข้อมูลดัชนีราคาผู้บริโภคของจังหวัดสระแก้วตั้งแต่ปี พ.ศ. 2545-2555 มาแบ่งหมวดหมู่ของดัชนีราคาผู้บริโภคออกเป็น 8 หมวดหมู่ จำนวน 80 แอททริบิวต์ โดยนำเทคนิคการจำแนก 4 แบบที่นิยมใช้มาเรียนรู้ข้อมูลสำหรับสร้างโมเดลพยากรณ์ โดยใช้โปรแกรม Orange Canvas เพื่อให้ได้กฎในการพยากรณ์ดัชนีราคาผู้บริโภค ซึ่งผลการทดสอบพบว่า อัลกอริทึม Rule Base (CN2) ให้ค่าความแม่นยำสูงสุด ได้กฎจำนวนทั้งสิ้น 16 กฎ และนำกฎที่ได้มาพัฒนาระบบพยากรณ์ดัชนีราคาผู้บริโภค จ.สระแก้ว เพื่อง่ายต่อการใช้งานระบบ และเพื่อง่ายต่อการสนับสนุนการตัดสินใจในการกำหนดราคาค่าแรง และค่าครองชีพของประชาชนจังหวัดสระแก้ว ซึ่งระบบดังกล่าวผู้จัดงานวิจัยได้นำกฎจากชุดข้อมูลดัชนีราคาผู้บริโภค จ.สระแก้ว มาทำการสร้างกฎ ดังนั้นกฎที่ได้ในการพัฒนาระบบฯ จึงมีความเหมาะสมสำหรับ จ.สระแก้ว เท่านั้น

เอกสารอ้างอิง

- [1] กิตติศักดิ์ อ่อนเอื้อน, พยุง มีสัง และ นวพร วิสิฐพงษ์พันธ์. “การวิเคราะห์หาปัจจัยที่ส่งผลกระทบต่อดัชนีราคาผู้บริโภคเพื่อรองรับเขตเศรษฐกิจพิเศษ ด้วยเทคนิคการทำเหมืองข้อมูล”. *NCCIT2013*. 2556. หน้า 96-101
- [2] สำนักงานดัชนีเศรษฐกิจการค้า สำนักงานปลัดกระทรวงพาณิชย์ กระทรวงพาณิชย์, “ความรู้เกี่ยวกับดัชนีเศรษฐกิจการค้า”, กรุงเทพฯ
- [3] สำนักงานจังหวัดสระแก้ว, “ข้อมูลจังหวัดสระแก้ว”. 2557
- [4] จังหวัดสระแก้ว, <http://www.sakaeco.go.th>, ค้นเมื่อวันที่ 26 ม.ค. 2557
- [5] สำนักงานพาณิชย์จังหวัดสระแก้ว, กระทรวงพาณิชย์, <http://pcoc.moc.go.th/>, ค้นเมื่อวันที่ 26 ม.ค. 2557
- [6] Ian H.Witten, Eibe Frank, Mark A. Hall. *Data Mining Practical Machine Learning Tool and Technique*, 3rd ed. Morgan Kaufmann: United States, 2011.
- [7] Jiao-Jiao Wang, Jin-Feng Wang, Feng Lu, Zhi-Dong Cao, Yi-Lan Liao and Yu Deng. “Comparison Study on Classification Performance for Short-Term Urban Traffic Flow Condition Using Decision Tree Algorithms”. *Software Engineering, 2009. WCSE '09. WRI World Congress*. 2009. pp.434-438
- [8] Zhu Xiaoliang, Wang Jian, Yan Hongcan and Wu Shangzhuo. “Research and application of the improved algorithm C4.5 on Decision tree”. *Test and Measurement, 2009. ICTM '09*. 2009. pp.184-187
- [9] Rafael S. Parpinelli, Heitor S. Lopes and Alex A. Freitas. “Data Mining With an Ant Colony Optimization Algorithm”. *IEEE Transaction on Evolutionary Computing*, vol. 6, no. 4, Aug 2002. pp.321
- [10] Pang-Ning Tan, Michael Steinbach and Vipin Kumar. *Introduction to Data Mining*. Boston San Francisco New York: Person Education, Inc. 2006.
- [11] Orange Canvas, <http://orange.biolab.si/features> ค้นเมื่อวันที่ 6 ก.พ. 2557.

ระบบการแนะนำในการเลือกซื้อโทรศัพท์มือถือด้วยกระบวนการกรองแบบพึ่งพาผู้ใช้ร่วม

Recommender System for Purchasing Mobile Phones using User-based Collaborative Filtering Approach

พนารัตน์ ศรีเชษฐา (Phanarut Srichetta)¹ และ พรพิมล พันธวาลี (Phonpimon Panthuvapee)²

สาขาวิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏอุดรธานี

¹phanarut@udru.ac.th, ²oumz@live.com

บทคัดย่อ

ในปัจจุบันนี้มีโทรศัพท์มือถือมากมายหลายยี่ห้อหลายรุ่น ถูกนำมาขายด้วยคุณสมบัติ คุณภาพ และราคาที่แตกต่างกัน การจะเลือกซื้อโทรศัพท์มือถือให้ตรงกับความต้องการของผู้ซื้อให้ได้มากที่สุดนั้น ผู้ซื้อต้องศึกษาและเปรียบเทียบ โทรศัพท์มือถือแต่ละยี่ห้อ/รุ่นจากแหล่งข้อมูลต่างๆ ด้วยตนเอง ซึ่งอาจใช้เวลานานกว่าจะได้ซื้อสรุป ดังนั้นงานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบให้คำแนะนำในการเลือกซื้อ โทรศัพท์มือถือที่เหมาะสมกับผู้ใช้แต่ละคน โดยใช้กระบวนการกรองแบบพึ่งพาผู้ใช้ร่วม (User-based Collaborative Filtering) ที่มีการคำนวณความคล้ายคลึงในการให้คะแนนความนิยมร่วมของกลุ่มผู้ใช้ในระบบกับผู้ใช้เป้าหมาย การคำนวณการทำนาย และการสร้างรายการโทรศัพท์มือถือแนะนำ โดยระบบทำงานแบบออนไลน์ผ่านทางอินเทอร์เน็ต สามารถช่วยผู้ใช้ในการค้นหารายละเอียดข้อมูลโทรศัพท์มือถือ ให้คะแนนความนิยมกับโทรศัพท์มือถือที่ตนเองชื่นชอบ และสามารถแนะนำโทรศัพท์มือถือที่เหมาะสมกับความชื่นชอบของผู้ใช้เป้าหมายได้

คำสำคัญ: การกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม การทำนายความคล้ายคลึง ระบบการแนะนำ โทรศัพท์มือถือ

Abstract

Nowadays many brands and models of mobile phones are sold with different property, quality, and price. In buying mobile phones to meet the needs of the buyer as much as possible, the buyers have to study and compare mobile phones in each brand and model from different sources by themselves. It may take longer to be concluded. Therefore, this research aims to develop a

system to advice on buying a mobile phone suitable for each person by using the user-based collaborative filtering approach. The system calculates the similarity of mobile phones' rating between the group of users in the system and the target user, calculates the prediction, and builds the list of recommended mobile phones by working online over the internet. It can help users to find the details of mobile phones, to rate their favorite mobile phones, and to recommend mobile phones that suit the preferences of the target user.

Keyword: User-based Collaborative Filtering, Prediction, Similarity, Recommender System, Mobile Phone.

1. บทนำ

เทคโนโลยีการสื่อสารในปัจจุบัน ถือได้ว่าเป็นปัจจัยสำคัญหนึ่งในชีวิตประจำวันของคนทุกระดับทุกสาขาอาชีพ ที่ใช้ในการติดต่อสื่อสารเพื่อให้ได้ข้อมูลที่สะดวกรวดเร็ว โดยเฉพาะเทคโนโลยีการสื่อสารแบบพกพาคือโทรศัพท์มือถือ ซึ่งในปัจจุบันมีการแข่งขันทางธุรกิจสูง มีการนำเสนอขายมากมายหลายยี่ห้อหลายรุ่น มีความสามารถในการทำงานที่หลากหลายมากขึ้นนอกเหนือจากการใช้พูดคุยสนทนาด้วยเสียง เช่น งานด้านความบันเทิง ด้านกราฟิก ด้านแผนที่หรือด้านเอกสารทั่วไป สำหรับการจะเลือกซื้อโทรศัพท์มือถือให้ได้ตามความต้องการและตอบสนองการใช้งาน ได้มากที่สุดนั้น ผู้ซื้อสามารถพิจารณาจากรายละเอียดต่างๆ ของโทรศัพท์มือถือที่ตนสนใจ เช่น ยี่ห้อ รุ่น รูปลักษณะ ราคา และประสิทธิภาพในการทำงานด้านต่างๆ จากแหล่งจำหน่ายหรือทางอินเทอร์เน็ต ต้องมีการเปรียบเทียบคุณลักษณะที่ใกล้เคียงกันของโทรศัพท์มือถือรุ่น

ต่างๆ เพื่อประกอบการตัดสินใจซื้อด้วยตนเอง ซึ่งทำให้ยุ่งยาก เสียเวลาในการหาข้อมูล รวมทั้งต้องตัดสินใจด้วยตัวเองโดยปราศจากเครื่องมือที่แนะนำรายการโทรศัพท์มือถือที่ตรงกับ ความชอบของผู้ซื้อมากที่สุด

ระบบให้คำแนะนำ (Recommender System) ถือได้ว่าเป็นเครื่องมือหนึ่งที่ช่วยแนะนำสินค้า/บริการ หรือรายการข้อมูลต่างๆ ให้แก่ผู้ใช้เพื่อแก้ปัญหาการตัดสินใจของผู้ใช้จากการที่มีข้อมูลมากเกินไป มีการนำระบบนี้ไปใช้อย่างกว้างขวางและกลายเป็นส่วนหนึ่งของการพาณิชย์อิเล็กทรอนิกส์ [1] เพื่อช่วยผู้ซื้อในการแนะนำสินค้าที่ตรงกับความต้องการได้อย่างมีประสิทธิภาพ ตัวอย่างเว็บไซต์ที่ประสบความสำเร็จในการนำระบบนี้ไปใช้ ได้แก่ ร้านขายหนังสือ Amazon [2] ร้านขายซีดี CDNow [3] และเว็บไซต์ที่ให้ข้อมูลเกี่ยวกับภาพยนตร์ MovieLens [4] เป็นต้น ดังนั้นระบบให้คำแนะนำนี้น่าจะเป็นเครื่องมือที่เหมาะสมต่อการนำมาใช้เพื่อช่วยแนะนำโทรศัพท์มือถือให้แก่ผู้ใช้ที่ต้องการเลือกซื้อ โดยทั่วไประบบให้คำแนะนำสามารถสร้างได้อาศัยคุณลักษณะของเนื้อหา รายการจากคำสำคัญต่างๆ ที่มีลักษณะคล้ายกัน หรืออาศัยความนิยมในรายการของผู้ใช้คนอื่นๆ ที่มีความนิยมคล้ายคลึงกับผู้ใช้ที่กำลังใช้งาน หรือใช้วิธีที่ผสมผสานระหว่างสองวิธีแรก [5]

การศึกษาในครั้งนี้ได้นำระบบให้คำแนะนำที่อาศัยความนิยมในรายการของผู้ใช้คนอื่นๆ ที่มีความนิยมคล้ายคลึงกับผู้ใช้ที่กำลังใช้งาน ที่เรียกว่า กระบวนการกรองแบบพึ่งพาผู้ใช้งาน (User-based Collaborative Filtering) มาใช้ในการสร้างรายการแนะนำการเลือกซื้อโทรศัพท์มือถือให้แก่ผู้ใช้ โดยระบบทำงานแบบออนไลน์ผ่านทางอินเทอร์เน็ต สามารถช่วยผู้ใช้ในการค้นหารายละเอียดข้อมูล โทรศัพท์มือถือ ให้คะแนนความนิยมกับโทรศัพท์มือถือที่ตนเองชื่นชอบได้ และสามารถแนะนำรายการโทรศัพท์มือถือที่เหมาะสมกับความนิยมชื่นชอบของผู้ใช้เป้าหมายที่กำลังใช้งานระบบมากที่สุด

2. กระบวนการกรองแบบพึ่งพาผู้ใช้งาน

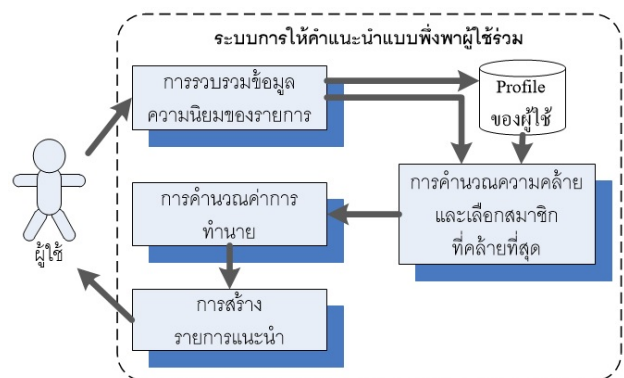
การกรองแบบพึ่งพาผู้ใช้งาน (Collaborative Filtering) เป็นวิธีการแนะนำรายการสินค้า/บริการ หรือรายการข้อมูลให้กับผู้ใช้ที่ใช้งานเป็นการส่วนบุคคล โดยอาศัยความนิยมของผู้ใช้คนอื่นๆ ที่มีรสนิยมหรือความชอบเหมือนกัน มีการประยุกต์ใช้

เทคนิคนี้ในหลายๆ ด้าน [7][8][9][10] โดยจุดเริ่มต้นของวิธีการนี้คือ ความนิยมของผู้ใช้ที่มีต่อรายการใดๆ ในระบบงานซึ่งสามารถป้อนเข้ามาได้โดยตรง เช่น ชอบ/ไม่ชอบ หรือเป็นค่าจำนวนจริงที่แสดงถึงระดับความชอบ คะแนนเหล่านี้จะถูกเก็บรวบรวมไว้เป็น Profile ของผู้ใช้แต่ละคนโดยภาพที่ 1 เป็นตัวอย่างการให้คะแนนความนิยมของผู้ใช้ในรูปของเมตริกซ์แถวอนคือผู้ใช้ แถวตั้งคือรายการ ค่า r_{ij} ในเมตริกซ์คือคะแนนความนิยมที่ผู้ใช้คนที่ i มีต่อรายการที่ j สัญลักษณ์ \bullet หมายถึงไม่มีการให้คะแนนนิยม โดยหากผู้ใช้เป้าหมายกับกลุ่มผู้ใช้คนอื่นๆ ได้ให้คะแนนความนิยมกับรายการร่วมกัน จะเรียกส่วนที่ให้คะแนนร่วมนั้นว่า โคเรท (Co-rate) เช่น ผู้ใช้ในกลุ่มคนที่ 1 และผู้ใช้เป้าหมายคนที่ u ได้ให้คะแนนร่วม (Co-rate) ในทุกรายการ ยกเว้นรายการที่ 2 และรายการที่ $n-1$

	1	2	...	i	...	$n-1$	N
1	$r_{1,1}$	$r_{1,2}$		$r_{1,i}$		\bullet	$r_{1,n}$
2	$r_{2,1}$			\bullet			
...							
u	$r_{u,1}$	\bullet		$r_{u,i}$		$r_{u,n-1}$	$r_{u,n}$

ภาพที่ 1: เมตริกซ์แสดงการให้คะแนนของผู้ใช้ร่วมแต่ละคน

การให้คะแนนรวมทั้งหมดของกลุ่มผู้ใช้ จะนำไปใช้ในการคำนวณความคล้ายคลึงกับ Profile ของผู้ใช้เป้าหมายที่กำลังใช้งาน จากนั้นจะใช้คะแนนความนิยมที่คล้ายคลึงมากที่สุด ในการทำนายคะแนนให้กับรายการที่ยังไม่ทราบคะแนน และสร้างรายการแนะนำให้ตรงกับความต้องการของผู้ใช้ที่กำลังใช้งานมากที่สุด ดังแสดงในภาพที่ 2



ภาพที่ 2: ระบบการให้คำแนะนำแบบพึ่งพาผู้ใช้งาน

2.1 การคำนวณความคล้ายคลึง

ในขั้นการคำนวณความคล้ายคลึงนั้น มีเทคนิคที่สามารถใช้ได้ 2 แบบ [5] คือ แบบอาศัยผู้ใช้ ที่เน้นการเปรียบเทียบโคเรลระหว่างผู้ใช้ (เปรียบเทียบแนวนอน) และแบบอาศัยรายการ ที่เน้นการเปรียบเทียบโคเรลระหว่างรายการ (เปรียบเทียบแนวตั้ง) โดยทั้งสองวิธีให้ผลลัพธ์เหมือนกัน ซึ่งในงานวิจัยนี้ผู้วิจัยเลือกใช้กระบวนการกรองแบบพึ่งพาร่วมโดยอาศัยผู้ใช้ (User-based Collaborative Filtering) สามารถทำการคำนวณความคล้ายคลึงระหว่างข้อมูล 2 ชุดได้โดยวิธี Correlation-based [6] ดังสมการที่ 1

$$sim(u_i, u_j) = \frac{\sum_{k=1}^N (r_{u_i,k} - \bar{r}_{u_i})(r_{u_j,k} - \bar{r}_{u_j})}{\sqrt{\sum_{k=1}^N (r_{u_i,k} - \bar{r}_{u_i})^2} \sqrt{\sum_{k=1}^N (r_{u_j,k} - \bar{r}_{u_j})^2}} \quad (1)$$

โดยที่

$r_{u_i,k}$ คือค่าคะแนนความนิยมที่ผู้ใช้ u_i มีต่อรายการที่ k

\bar{r}_{u_i} คือค่าคะแนนเฉลี่ยของรายการร่วมที่ให้โดยผู้ใช้ u_i

N คือจำนวนของรายการที่ให้คะแนนร่วม (co-rated items)

ตารางที่ 1 เป็นตัวอย่างที่มีผู้ใช้ 4 คน (u_1, u_2, u_3, u_4) ให้คะแนนนิยมนับรายการสินค้าในระบบ 4 รายการ ซึ่งจะแน่นอนอยู่ในช่วง 1-5 โดย 1 หมายถึงมีความนิยมน้อยมาก 5 หมายถึงมีความนิยมมาก * หมายถึงไม่ได้ให้คะแนนความนิยม โดยกำหนดให้ u_i เป็นผู้ใช้ที่กำลังใช้งาน ภายในตารางนี้ยังแสดงคะแนนเฉลี่ยของผู้ใช้แต่ละคนตามการให้คะแนนร่วม

ตารางที่ 1: ตัวอย่างการให้คะแนนนิยมนับรายการและค่าเฉลี่ย

	1	2	3	4	คะแนนเฉลี่ยตามผู้ใช้ (\bar{r}_{u_i})
u_1	2	1	3	5	$S(u_1) = (2+3+5)/3 = 3.3333$
u_2	*	2	1	4	$S(u_2) = (1+4)/2 = 2.5$
u_3	3	2	*	4	$S(u_3) = (3+4)/2 = 3.5$
u_4	4	3	5	*	$S(u_4) = (4+5)/2 = 4.5$
u_i	4	*	5	5	$S(u_i) = (4+5+5)/3 = 4.6667$

ค่าความคล้ายคลึงระหว่างผู้ใช้ u_i กับผู้ใช้คนอื่นที่ 1 สามารถแสดงด้วยตัวอย่างด้านล่างนี้ และความคล้ายคลึงกับผู้ใช้คนอื่นๆ แสดงดังตารางที่ 2

$$sim(u_i, u_i) = \frac{[(4-4.67)(2-3.33) + (5-4.67)(3-3.33) + (5-4.67)(5-3.33)]}{\sqrt{(4-4.67)^2 + (5-4.67)^2 + (5-4.67)^2} \sqrt{(2-3.33)^2 + (3-3.33)^2 + (5-3.33)^2}} = \frac{1.33}{\sqrt{0.67} \sqrt{0.67}} = 0.7559$$

ตารางที่ 2: ตัวอย่างผลลัพธ์การคำนวณค่าความคล้ายคลึง

$sim(u_i, u_1)$	$sim(u_i, u_2)$	$sim(u_i, u_3)$	$sim(u_i, u_4)$
0.7559	0	0.9486	0.9486

ค่าความคล้ายคลึงที่ได้จะมีค่าอยู่ในช่วง -1 ถึง 1 โดยค่าลบนั้นไม่มีค่าใดๆ ในการเพิ่มประสิทธิภาพในการทำนาย สามารถเลือกใช้โดยไม่ต้องพิจารณาที่เครื่องหมายลบ โดยจะใช้ค่าความคล้ายคลึงไปหาเซตของเพื่อนบ้านที่ใกล้เคียงที่สุด (Nearest neighbors) จากการกำหนดค่าขีดแบ่ง (threshold: δ) หรือการกำหนดจำนวนรายการสูงสุด N รายการ (Top-N) นั่นคือ เซตของผู้ใช้ที่คล้ายคลึงกับผู้ใช้ที่กำลังใช้งาน u_i มากที่สุด ($S(u_i)$)

$$S(u_i) = \{u_j \mid sim(u_i, u_j) \geq \delta\} \quad (2)$$

$$\text{หรือ } S(u_i) = \{u_j \mid rank sim(u_i, u_j) \leq N\} \quad (3)$$

ดังนั้นหากระบุ $\delta = 0.8$ จะได้ $S(u_i) = \{u_3, u_4\}$ หรือหากกำหนดจำนวนรายการสูงสุด 2 รายการ (Top-2) จะได้ $S(u_i) = \{u_3, u_4\}$

2.2 การคำนวณค่าการทำนาย (Prediction Computation)

ขั้นตอนี่ใช้ในการคำนวณหาค่าคะแนนที่คาดว่าผู้ใช้ที่กำลังใช้งาน u_i จะมีต่อรายการ k ที่ยังไม่ได้ให้คะแนนไว้ ($P_{u_i,k}$) โดยใช้ค่าความคล้ายคลึงและค่าคะแนนนิยมนับร่วมจากกลุ่มของผู้ใช้ที่คล้ายคลึงกับผู้ใช้ที่กำลังใช้งานมากที่สุด $S(u_i)$

$$P_{u_i,k} = \bar{r}_{u_i} + \frac{\sum_{u_j \in S(u_i)} sim(u_i, u_j) \cdot (r_{u_j,k} - \bar{r}_{u_j})}{\sum_{u_j \in S(u_i)} sim(u_i, u_j)} \quad (4)$$

โดยที่

\bar{r}_{u_i} คือค่าเฉลี่ยคะแนนนิยมนับของผู้ใช้ที่กำลังใช้งาน u_i

$sim(u_i, u_j)$ คือค่าความคล้ายคลึงระหว่างผู้ใช้ที่กำลังใช้งาน u_i และผู้ใช้ u_j ใดๆในเซตของผู้ใช้ที่คล้ายคลึงกับ u_i ที่สุด

$r_{u_j,k}$ คือค่าคะแนนความนิยมที่ผู้ใช้ u_j มีต่อรายการ k

$S(u_i)$ คือเซตของผู้ใช้ใดๆที่คล้ายกับผู้ใช้ u_i มากที่สุด

จากตัวอย่างก่อนหน้า ผู้ใช้ที่กำลังใช้งาน u_i ยังไม่ได้ให้คะแนนนิยมนับรายการที่ 2 สามารถคำนวณค่าการทำนายโดยอาศัยกลุ่มของสมาชิกที่คล้ายคลึงที่สุดจาก Top-2 ซึ่งก็คือ $S(u_i) = \{u_3, u_4\}$ ได้ค่าการทำนายเป็น

$$P_{u_i,2} = \bar{r}_{u_i} + \frac{sim(u_i, u_3)(r_{u_3,2} - \bar{r}_{u_3}) + sim(u_i, u_4)(r_{u_4,2} - \bar{r}_{u_4})}{sim(u_i, u_3) + sim(u_i, u_4)}$$

$$= 4.6667 + \frac{(0.9486)(2 - 3.5) + (0.9486)(3 - 4.5)}{0.9486 + 0.9486} = 3.1667$$

โดยการจะสร้างรายการแนะนำให้กับผู้ใช้ u_i นั้น ทุกรายการต้องทำการคำนวณค่าการทำนายในทำนองเดียวกับรายการที่ 2 ผลการคำนวณค่าการทำนายให้กับผู้ใช้ u_i แสดงไว้ในตารางที่ 3

ตารางที่ 3: ตัวอย่างผลลัพธ์การคำนวณค่าการทำนาย

$P_{u_i,1}$	$P_{u_i,2}$	$P_{u_i,3}$	$P_{u_i,4}$
4.1667	3.1667	4.7467	4.9167

2.3 การให้คำแนะนำ (Recommendation)

ขั้นตอนนี้เป็นกรนำค่าการทำนายแต่ละรายการมาเรียงลำดับจากมากไปน้อย โดยเลือกรายการแนะนำให้กับผู้ใช้ที่กำลังใช้งาน N รายการแรก (Top-N) ขึ้นอยู่กับผู้สร้างระบบว่าจะให้แสดงกี่รายการ ดังนั้น รายการแนะนำให้กับผู้ใช้ u_i คือ

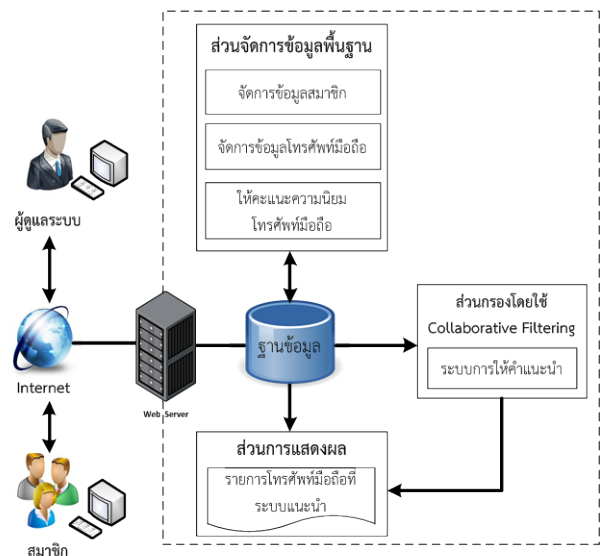
$$R(u_i) = \{k | rank(P_{u_i,k}) \leq N\} \tag{5}$$

ดังนั้นจากตัวอย่าง หากต้องการแนะนำให้กับผู้ใช้ u_i เพียง 2 รายการแรก (Top-2) จากรายการทั้งหมดที่มีการให้คะแนนรวมจะได้สินค้ารายการที่ 4 เป็นลำดับที่ 1 ($P_{u_i,4} = 4.9167$) และสินค้ารายการที่ 3 เป็นลำดับที่ 2 ($P_{u_i,3} = 4.7467$)

3. วิธีการดำเนินการศึกษา

สำหรับข้อมูลที่ใช้ในการศึกษาครั้งนี้ รวบรวมข้อมูลโทรศัพท์มือถือจากเว็บไซต์ผู้ผลิต ได้แก่ บริษัทโนเกีย บริษัทซัมซุง บริษัทอริคสัน เป็นต้น นอกจากนี้จากเว็บไซต์ที่รวบรวมข้อมูลของโทรศัพท์มือถือ www.siamphone.com สามารถแบ่งยี่ห้อของโทรศัพท์มือถือที่ใช้เป็นกรณีศึกษาได้ 9 ยี่ห้อคือ Nokia, HTC, LG, Samsung, Motorola, i-Mobile, BlackBerry, G-net และ Sony Ericsson แต่ละยี่ห้อยังมีได้หลายรุ่นย่อยลงไปที่คุณสมบัติและราคาแตกต่างกันไป ซึ่งข้อมูลโทรศัพท์มือถือที่ใช้ในการศึกษาครั้งนี้เป็นโทรศัพท์ที่ผลิตและวางจำหน่าย 5 ปีย้อนหลังจำนวน 350 รุ่นจาก 9 ยี่ห้อ โทรศัพท์แต่ละรายการประกอบด้วยข้อมูลยี่ห้อ รุ่น ราคา รูปประกอบ และคุณสมบัติฟังก์ชันการใช้งานในด้านต่างๆ เช่น ฟังเพลง mp3 วิทยู อินฟราเรด บลูทูท เป็นต้น

แนวคิดของระบบการแนะนำในการเลือกซื้อโทรศัพท์มือถือด้วยกระบวนการกรองแบบพึ่งพาผู้ใช้ร่วมกัน จะแบ่งการทำงานออกเป็น 3 ส่วนหลักคือ 1) ส่วนจัดการข้อมูลพื้นฐาน ซึ่งใช้จัดการการเพิ่ม ลบ แก้ไข และค้นหาข้อมูลพื้นฐานต่างๆ ของผู้ใช้สมาชิก โทรศัพท์มือถือ และคะแนนความนิยมของผู้ใช้ที่มีต่อโทรศัพท์มือถือ จัดเก็บไว้ในฐานข้อมูล 2) ส่วนการกรองโดยใช้ Collaborative Filtering ซึ่งจะเป็นการประมวลผลรายการแนะนำโทรศัพท์มือถือตามเทคนิคที่ได้นำเสนอโดยใช้ข้อมูลจากฐานข้อมูล และนำผลไปเสนอใน 3) ส่วนแสดงผล ดังภาพรวมระบบที่แสดงในภาพที่ 3



ภาพที่ 3: ภาพรวมระบบแนะนำโทรศัพท์มือถือ

ระบบการแนะนำในการเลือกซื้อโทรศัพท์มือถือที่พัฒนาเป็นเว็บแอปพลิเคชันทำงานผ่านเครือข่ายอินเทอร์เน็ต โดยผู้ใช้ระบบประกอบด้วยผู้ดูแลระบบที่มีหน้าที่ในการจัดการข้อมูลพื้นฐานโทรศัพท์มือถือ และผู้ใช้ทั่วไปที่สามารถสมัครเป็นสมาชิกเพื่อรับข้อมูลเกี่ยวกับโทรศัพท์มือถือ โดยระบบนี้พัฒนาด้วยภาษา PHP และมี MySQL เป็นระบบการจัดการฐานข้อมูล

4. ผลการดำเนินงาน

จากการศึกษาและพัฒนากระบวนการแนะนำในการเลือกซื้อโทรศัพท์มือถือด้วยกระบวนการกรองแบบพึ่งพาผู้ใช้ร่วมกัน มีผลการดำเนินการดังนี้

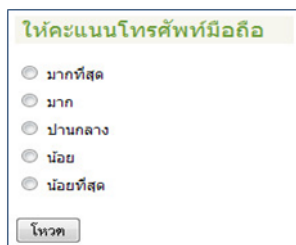
4.1 ส่วนการจัดการข้อมูลพื้นฐาน

ในส่วนการจัดการข้อมูลพื้นฐาน โทรศัพท์มือถือนั้นเป็นหน้าที่ของผู้ดูแลระบบ ที่จะเป็นผู้เพิ่ม ลบ หรือแก้ไขข้อมูล ดังแสดงในภาพที่ 4 โดยการค้นหารายละเอียดของโทรศัพท์มือถือสามารถทำได้โดยผู้ดูแลระบบ สมาชิก และผู้ใช้ทั่วไป



ภาพที่ 4: หน้าจอส่วนการจัดการข้อมูลโทรศัพท์มือถือ

สำหรับการจัดการข้อมูลสมาชิกนั้น เป็นหน้าที่ของผู้ใช้ทั่วไปที่สามารถสมัครเข้าเป็นสมาชิกของระบบก่อน จึงจะมีสิทธิ์ในการแก้ไขข้อมูลของตนเอง รวมทั้งแสดงความคิดเห็น และสามารถโหวตให้คะแนนโทรศัพท์มือถือแต่ละยี่ห้อ/รุ่น ที่ตนเองได้พิจารณารายละเอียดข้อมูลโทรศัพท์มือถือที่มีในระบบแล้ว โดยไม่สามารถโหวตซ้ำได้ เนื่องจากระบบทำการบันทึกค่าการโหวตล่าสุด ซึ่งผู้ใช้แต่ละคนสามารถโหวตได้หลายรุ่น โดยการเลือกรุ่นที่ชอบมากที่สุดแล้วรอกคะแนนโหวตให้รุ่นนั้นเพื่อเก็บลงในโพรไฟล์ของผู้ใช้แต่ละคนในฐานข้อมูล ดังภาพที่ 5



ภาพที่ 5: หน้าจอส่วนให้คะแนนความนิยมโทรศัพท์มือถือ

4.2 ส่วนการกรองโดยใช้ Collaborative Filtering

ในส่วนการกรองโดยใช้ Collaborative Filtering นี้เป็นการทำงานโดยอัตโนมัติ นั่นคือทุกครั้งที่มีการให้คะแนนความ

นิยมไม่ว่าจากผู้ใช้คนใด ระบบจะประมวลผลการคำนวณความคล้ายคลึงจากการให้คะแนนโหวตร่วมระหว่างผู้ใช้ในระบบฐานข้อมูลกับผู้ใช้เป้าหมายที่กำลังใช้งานระบบอยู่ ซึ่งค่า threshold (δ) ที่ใช้ระบบคือ 0.8 เพื่อเลือกเฉพาะกลุ่มที่เกือบคล้ายคลึงมากที่สุด จากนั้นการคำนวณค่าการทำนาย และการเตรียมรายการโทรศัพท์แนะนำ สำหรับผู้ใช้เป้าหมายที่กำลังใช้งานอยู่ ณ ขณะนั้น โดยในขั้นนี้ ต้องอยู่ภายใต้สมมติฐานที่ว่าผู้ใช้ในระบบควรมีมากพอ รวมทั้งผู้ใช้ต้องมีการโหวตให้คะแนนความชอบกับโทรศัพท์มือถือที่มีในระบบเป็นจำนวนมาก จึงจะทำให้การสร้างรายการแนะนำโทรศัพท์มือถือทำได้ อย่างมีประสิทธิภาพ

4.3 ส่วนการแสดงผล

สำหรับส่วนการแสดงผลนี้ เป็นการสร้างรายการแนะนำโทรศัพท์มือถือที่ใกล้เคียงกับที่ผู้ใช้เป้าหมายที่กำลังใช้งานระบบอยู่ต้องการ เมื่อเทียบกับการโหวตร่วมกับผู้ใช้คนอื่นๆ ในระบบ ดังภาพที่ 6



ภาพที่ 6: หน้าจอส่วนแสดงผลการให้คำแนะนำโทรศัพท์มือถือ

4.4 การประเมินผลระบบที่พัฒนา

ระบบที่พัฒนานี้ ได้นำไปทดลองกับผู้ใช้ทั่วไปจำนวน 20 คน ที่ประสงค์จะเลือกซื้อ โทรศัพท์มือถือและที่เปลี่ยนโทรศัพท์มือถือเป็นประจำตามวิวัฒนาการของเทคโนโลยีที่เปลี่ยนไป ทำให้มีความชอบ/ไม่ชอบโทรศัพท์มือถือในหลากหลายยี่ห้อ/รุ่น เพื่อประเมินประสิทธิภาพการทำงานของระบบ โดยรายการประเมินในด้านต่างๆ ที่มีคะแนนเต็ม 5 และผลการประเมินเป็นดังนี้

- ด้านความครอบคลุมของระบบงานที่พัฒนา กับระบบงานจริง ($\bar{x} = 4.19, S.D.=0.63$)

- ด้านความเหมาะสมในการโต้ตอบกับผู้ใช้ (ความง่ายในการใช้งาน/ ความเป็นมาตรฐานเดียวกันของหน้าจอ/การมีส่วนช่วยเหลือ) ($\bar{x} = 4.68$, S.D.=0.41)
- ด้านความถูกต้องของรายการแนะนำโทรศัพท์ตรงกับความต้องการผู้ใช้ ($\bar{x} = 4.04$, S.D.=0.81)
- ด้านความรวดเร็วในการประมวลผลของระบบ ($\bar{x} = 4.92$, S.D.=0.27)
- ด้านความน่าเชื่อถือของระบบ ($\bar{x} = 4.39$, S.D.=0.62)

โดยประเด็นการประเมินในด้านความถูกต้องของรายการแนะนำโทรศัพท์มือถือว่าตรงกับความต้องการผู้ใช้หรือไม่ได้นั้น ได้ค่าเฉลี่ยเป็น 4.36 และค่าส่วนเบี่ยงเบนมาตรฐานเป็น 0.81 ซึ่งถือว่ามี การสร้างรายการแนะนำอยู่ในระดับดี และผลการประเมินโดยรวมในทุกด้าน ค่าเฉลี่ยเป็น 4.67 ค่าเบี่ยงเบนมาตรฐาน 0.66 ซึ่งถือว่าโดยรวมระบบสามารถทำงานมีประสิทธิภาพอยู่ในระดับดี

5. บทสรุป

การศึกษาในครั้งนี้ ผู้วิจัยได้พัฒนาระบบการแนะนำในการเลือกซื้อโทรศัพท์มือถือแบบออนไลน์ ทำงานผ่านเครือข่ายอินเทอร์เน็ต โดยระบบมีส่วนการจัดการข้อมูลพื้นฐานเกี่ยวกับสมาชิก โทรศัพท์มือถือ ความคิดเห็นเกี่ยวกับโทรศัพท์มือถือ และการให้คะแนนความนิยมกับโทรศัพท์มือถือแต่ละยี่ห้อ/รุ่น สำหรับการกรองโดยใช้ Collaborative Filtering จะมีการคำนวณหาความคล้ายคลึงระหว่างความนิยมโทรศัพท์มือถือของผู้ใช้เป้าหมายกับผู้ใช้ในระบบ คำนวณค่าการทำนายโทรศัพท์มือถือที่ผู้ใช้เป้าหมายยังไม่ได้ให้คะแนนความนิยม และการคำนวณสร้างรายการแนะนำโทรศัพท์มือถือ ให้กับผู้ใช้เป้าหมายที่กำลังใช้งานอยู่เป็นการส่วนบุคคลให้ตรงตามความนิยมของผู้ใช้มากที่สุด โดยจะแสดงผลรายการแนะนำในส่วนแสดงผล ซึ่งผลการประเมินประสิทธิภาพการทำงานของระบบ โดยผู้ใช้อาจจะระบบทำงานอยู่ในระดับดี

ในการที่จะนำระบบนี้ไปใช้งานให้มีประสิทธิภาพเพื่อสร้างรายการแนะนำโทรศัพท์มือถือให้กับผู้ใช้เป้าหมายได้อย่างถูกต้องนั้น จำเป็นต้องมีผู้ใช้ในระบบจำนวนมาก รวมทั้งผู้ใช้

ต้องมีการให้คะแนนความนิยมกับโทรศัพท์มือถือแต่ละยี่ห้อ/รุ่น ที่มีอยู่ในระบบเป็นจำนวนมาก

เอกสารอ้างอิง

- [1] A. Ansari, S. Essegaier and R. Kohli, "Internet Recommendation System", *Journal of Marketing Research*, vol. 37, no. 3, pp. 363-375, 2000.
- [2] G. Linden, B. Smith and J. York, "Amazon.com Recommendation: Item-to-Item Collaborative Filtering", *IEEE Internet Computing*, vol. 7, no. 1, pp. 76-80, 2003.
- [3] CDNNow data, <http://cdnow.com/>
- [4] MovieLen data, <http://www.grouplens.org/>
- [5] G. Adomavicius and A. Tuzhilin, "Toward the Next Generation of Recommender System: A Survey of the State-of-the-Art and Possible Extensions", *IEEE Transaction on Knowledge and Data Engineering*, vol. 17, no. 6, pp. 734-740, 2005.
- [6] M.R. McLaughlin and J.L. Herlocker, "A Collaborative Filtering Algorithm and Evaluation Metric that Accurately Model the User Experience", *Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 329-336, 2004.
- [7] ทศนวรรณ แก้วใส และสุพจน์ นิตย์สุวรรณ. "ระบบแนะนำภาพยนตร์ด้วยเทคนิค Collaborative Filtering ร่วมกับ K-Mean" ปัญหาพิเศษปริญญาวิทยาศาสตรมหาบัณฑิตสาขาเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ. 2551.
- [8] สุภาวดี สุกถ้อย และจิรารัตน์ สิทธิวรชาติ. "ระบบแนะนำเพลงสำหรับการดาวน์โหลดผ่านโทรศัพท์มือถือด้วยเทคนิค Collaborative Filtering" ปัญหาพิเศษปริญญาวิทยาศาสตรมหาบัณฑิต สาขาเทคโนโลยีสารสนเทศ มหาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ. 2551.
- [9] ปรีดา สมสังข์ และมณเฑียร รัตนศิริวงศ์วุฒิ. "การพัฒนา ระบบแนะนำรายวิชาเรียนโดยใช้เทคนิค Collaborative Filtering" ปัญหาพิเศษปริญญาวิทยาศาสตรมหาบัณฑิต สาขาการจัดการเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ. 2551.
- [10] สุภรัตน์ รัชมีเทศ. "ระบบให้คำแนะนำในการเลือกซื้อเลนส์สำหรับกล้องดีเอสแอลอาร์ด้วยเทคนิค Content-based Filtering และ Collaborative Filtering" ภาควิชาเทคโนโลยีสารสนเทศมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ. 2554.

ระบบแนะนำสถานที่ท่องเที่ยวอุทยานแห่งชาติทางภาคใต้ของประเทศไทยด้วย กูเกิลแมพ

เอพีไอ บนมือถือ สมาร์ทโฟน

Recommender Systems for South National Park of Thailand Tourism Using Google Maps API on Smart Phone

ชไมพร ทองขาว (Chamaiporn Tongkaw)¹ และ มาลีรัตน์ โสคานิล (Maleerat Sodanil)²

^{1,2}ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
comscience28@hotmail.com¹, msn@kmutnb.ac.th²

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบแนะนำสถานที่ท่องเที่ยวอุทยานแห่งชาติทางภาคใต้ของประเทศไทย โดยใช้กูเกิลแมพ และใช้เทคนิค Collaborative Filtering ร่วมกับ K-Mean มาช่วยในการแนะนำข้อมูลจากชั้นข้อมูลเรตติ้ง (Rating) ในระบบเทียบกับเรตติ้งของผู้ใช้งาน และทำนายสถานที่ที่คาดว่าตรงกับความต้องการของผู้ใช้ โดยใช้อัลกอริทึม K-Mean ในการหาค่าที่มีความใกล้เคียงของข้อมูล และประเมินความพึงพอใจในการใช้ระบบโดยการหาค่าเฉลี่ย (Mean) และค่าส่วนเบี่ยงเบนมาตรฐาน (S.D) โดยกลุ่มประชากรที่ทำการทดสอบเป็นกลุ่มประชากร ในกรุงเทพฯ จำนวน 20 คน และทดสอบประสิทธิภาพ โดยผู้เชี่ยวชาญ จำนวน 3 ท่าน ผลการประเมินความพึงพอใจของผู้ใช้ระบบ ($\bar{x}=4.50$, S.D = 0.02) และผลการประเมินประสิทธิภาพโดยผู้เชี่ยวชาญ ($\bar{x}=4.38$, S.D = 0.06)

คำสำคัญ: อุทยานแห่งชาติทางภาคใต้ ระบบแนะนำ
ตัวกรองเชิงร่วมมือ

Abstract

This research aims to develop a system that can recommend south national park of Thailand tourist attraction. By using google map for navigation. And using collaborative filter with k-mean technic for recommendation by comparison system rating and users rating to predicts tourist attraction where relevance need of the user. By using k-mean algorithm for finding a point that nearly a data.

System usage satisfaction evaluation by 20 samples of Bangkok tourist and system performance evaluation by 3 experts. The results of system usage satisfaction ($\bar{x}=4.50$, S.D = 0.02) and the result of performance evaluation by specialist ($\bar{x}=4.38$, S.D = 0.06)

Keywords : South National Park Recommender Systems Collaborative Filtering

1. บทนำ

สถานที่ท่องเที่ยวทางอุทยานแห่งชาติของภาคใต้นั้นมีความสวยงามและเป็นสถานที่ที่นักท่องเที่ยวต่างให้ความสนใจในปัจจุบันอุทยานแห่งชาตินั้นได้มีการก่อตั้งขึ้นมาด้วยกันทั้งหมด 40 อุทยานแห่งชาติ แบ่งเป็น 2 กลุ่มใหญ่ คือ อุทยานแห่งชาติทางบก 22 แห่ง และอุทยานแห่งชาติทางทะเล 18 แห่ง [1] และยังมีอีกบางสถานที่ที่กำลังจะได้รับการแต่งตั้งเป็นอุทยานแห่งชาติ

จากการที่นักท่องเที่ยวให้ความสนใจกับสถานที่ท่องเที่ยวทางภาคใต้มากขึ้นเรื่อย ๆ และสถานที่ท่องเที่ยวเกิดขึ้นใหม่เพิ่มมากขึ้นทำให้นักท่องเที่ยวใหม่ๆที่เข้ามาท่องเที่ยวทางภาคใต้นั้นไม่รู้จักกับสถานที่ท่องเที่ยวที่หลากหลายผู้วิจัยจึงได้เก็บรวบรวมความต้องการของนักท่องเที่ยวว่ามีนักท่องเที่ยวที่มีความต้องการในด้านใดบ้างและสรุปได้ว่าเราควรจะมีระบบแนะนำให้แก่นักท่องเที่ยวเพื่ออำนวยความสะดวกให้แก่นักท่องเที่ยวที่มาท่องเที่ยวอุทยานแห่งชาติทางภาคใต้

ดังนั้นผู้วิจัยจึงได้สร้างระบบแนะนำสถานที่ท่องเที่ยวอุทยานแห่งชาติทางภาคใต้ขึ้น โดยพัฒนาบนระบบปฏิบัติการ

แอนดรอยด์และใช้ กูเกิลแมพ เอพีไอ มาช่วยในการนำทางให้กับนักท่องเที่ยวและใช้เทคนิค Collaborative Filtering มาช่วยหาความสัมพันธ์ของข้อมูลเพื่อใช้ในการจัดเรตติ้งให้กับสถานที่ท่องเที่ยว เพื่อให้สถานที่ท่องเที่ยวที่มีความน่าสนใจเพิ่มมากขึ้น

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ระบบแนะนำ(Recommender Systems)

เป็นการนำเสนอประสบการณ์กับบุคคลให้เกิดขึ้นกับผู้ใช้ ด้วยการนำเสนอข้อมูลที่ถูกล้นกรองสำหรับผู้ใช้งานแต่ละคน ผ่านช่องทางการติดต่อสื่อสารเป็นการเพิ่มโอกาสให้ผู้ใช้งานได้รับข้อมูลที่ตรงตามความต้องการมากที่สุด ในงานวิจัยนี้ได้ นำระบบแนะนำมาช่วยแก้ปัญหาจำนวนชิ้นข้อมูลที่มีจำนวนมากที่ยังไม่ได้คัดกรองนำมาคัดกรองให้เหมาะสมกับผู้ใช้งานมากที่สุด ชิ้นข้อมูลแต่ละชิ้นจะถูกแนะนำขึ้นอยู่กับตัวชี้วัด นั่นคือค่าเรตติ้ง [2]

2.2 ระบบปฏิบัติการแอนดรอยด์ (Android)

เป็นชื่อเรียกซอฟต์แวร์หรือแพลตฟอร์มสำหรับอุปกรณ์อิเล็กทรอนิกส์ที่มีหน่วยประมวลผล [3] เป็นส่วนประกอบ เช่น โทรศัพท์เคลื่อนที่ สมาร์ทโฟน คอมพิวเตอร์ และอุปกรณ์พกพา เป็นการนำเอาระบบปฏิบัติการลินุกซ์ (Linux) ซึ่งนำไปใช้งานกับเครื่องเซิร์ฟเวอร์ (Server) เป็นหลัก ที่มีขนาดพื้นที่จัดเก็บข้อมูลที่จำกัดโดยที่แอนดรอยด์จะใช้สัญลักษณ์หุ่นยนต์ เป็นหุ่นยนต์ที่คอยช่วยเหลืออำนวยความสะดวกแก่ผู้ที่พกพาในทุกที่ทุกเวลา ระบบปฏิบัติการแอนดรอยด์สามารถเชื่อมต่อกับ API(Application Program Interface) การเรียกใช้ระบบปฏิบัติการหรือแอปพลิเคชันอื่นๆ หรือชุดโค้ด คอมพิวเตอร์ซึ่งทำหน้าที่เชื่อมต่อการทำงานระหว่างแอปพลิเคชันกับระบบปฏิบัติการซึ่งในงานวิจัยนี้มีการนำเอพีไอมาใช้ 2 ชนิด คือ

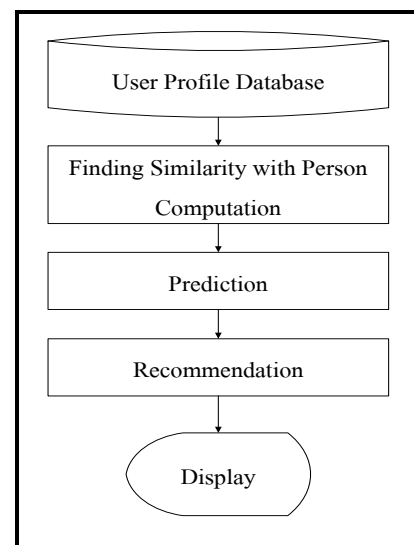
กูเกิลแมพ เอพีไอ เป็น Web Service API เป็นบริการของกูเกิลแบ่งออกเป็น 2 ส่วน คือ Web Service Server บริการกูเกิลแมพแสดงเป็นข้อมูลแผนที่ต่างๆ และ Web Service Client เป็นโปรแกรมที่สร้างมาเพื่อเรียกใช้งาน Web Service เป็นหนึ่งกูเกิลเว็บที่ได้รับความนิยมมากในปัจจุบันนักพัฒนาสามารถพัฒนาแผนที่ที่มีความละเอียดสูงในการประยุกต์ใช้ต่าง ๆ เช่น การจราจร การโปรโมทสินค้า การสำรวจได้ง่าย [4]

เฟซบุ๊ก เอพีไอ เป็นบริการจากเฟซบุ๊ก ที่ช่วยอำนวยความสะดวกให้กับผู้พัฒนาในการที่จะนำข้อมูลที่อยู่ในเฟซบุ๊กไปใช้ เช่น ข้อมูลการไลค์ ข้อมูลการเช็คอิน ข้อมูลรูปภาพ ข้อมูลการคอมเม้น และอีกหลายฟังก์ชันที่สามารถนำมาใช้ได้ Graph API ซึ่งตัวมันเป็น REST API (get, post, put, delete) โดยตัวนี้เป็นตัวจัดการข้อมูล Facebook โดยทางเราต้องส่ง URL ที่ต้องการแล้วพอมันทำงานเสร็จจะส่งข้อมูลกลับมาในรูปแบบ JSON ทำให้เราสามารถดึงข้อมูลมาใช้งานใน Application ของเราได้ อย่างเต็มที่

2.3 การกรองเชิงร่วมมือ(Collaborative Filtering)

เทคนิคการกรองข้อมูลแบบพึ่งพา (CF) คือวิธีที่ใช้เพื่อทำนายหรือคัดกรองสิ่งที่ผู้ใช้สนใจด้วยข้อมูลรสนิยมซึ่งเก็บรวบรวมจากผู้ใช้งานหลายคน ด้วยสมมุติฐานที่ว่าผู้ที่มีความสนใจหรือรสนิยมในอดีตเหมือนกันน่าจะมีความสนใจเหมือนกันในอนาคตเช่นกัน เนื่องจากในอนาคตข้อมูลจะมีปริมาณเพิ่มขึ้นซึ่งทำให้ผู้ใช้งานไม่สามารถพิจารณาของทุก ๆ สิ่งได้ในเวลาที่จำกัดทำให้เทคนิคการกรองข้อมูลแบบพึ่งพา ถูกประยุกต์ใช้เพื่อสร้างระบบให้คำแนะนำ เพื่อแนะนำสิ่งของให้แก่ผู้ใช้งาน โดยเป็นสิ่งของที่มีประโยชน์ต่อผู้ใช้งาน หรือผู้ใช้งานจะมีความชื่นชอบในสิ่งของที่ถูกแนะนำนี้

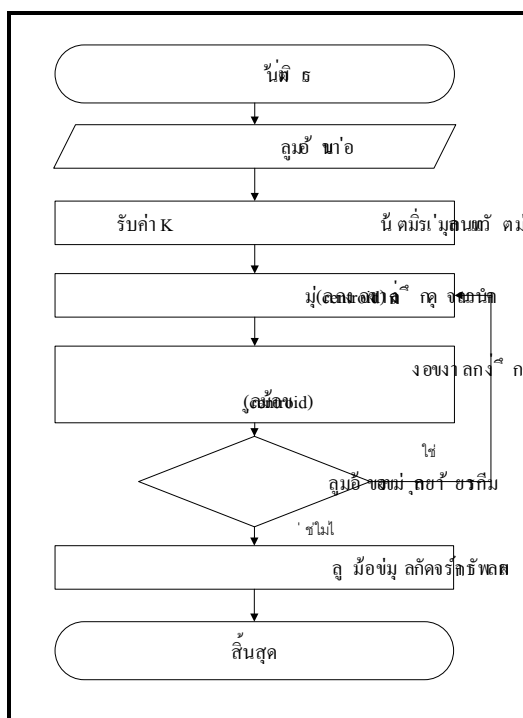
กระบวนการทำงานประกอบด้วย 3 ขั้นตอนคือ การหาค่าความคล้ายคลึงของข้อมูล (Similarity Computation) การทำนาย (Prediction) และการสร้างรายการแนะนำ (Recommendation)



ภาพที่ 1: กระบวนการทำงานของ Collaborative Filtering

2.4 เคมีน (K-Mean)

เป็นวิธีที่ง่ายทำงานรวดเร็ว และนิยมใช้มากใช้ในการแบ่งกลุ่มแบบตัดเป็นส่วน และวัดคุณภาพโดยค่าความคลาดเคลื่อนยกกำลังสองส่วน จะเหมาะสำหรับการแบ่งกลุ่มข้อมูลที่มีข้อมูลแต่ละกลุ่มมีขนาดใกล้เคียงกัน ในทางปฏิบัติสำหรับการจัดการกับตัวแปรอิสระที่มีมากกว่า 2 ตัวแปร หมายความว่าใช้ตำแหน่งเหมือนการแทนที่ 2 องค์ประกอบ Vectors(x_1, x_2) ด้วยตำแหน่งที่เหมือนกันถึง n องค์ประกอบ Vectors(x_1, x_2, \dots, x_n) โดยจะกำหนดจำนวนกลุ่มไว้ล่วงหน้ามีขั้นตอนการทำงานเริ่มจากการสุ่มเลือกจุดศูนย์กลางของแต่ละกลุ่ม ต่อมาวัดระยะห่างระหว่างข้อมูลแต่ละตัวกับจุดศูนย์กลางกลุ่ม โดยอาศัยเงื่อนไขของค่าที่ใกล้ที่สุดหรือระยะห่างน้อยที่สุดในการแบ่งกลุ่มข้อมูลจากนั้นคำนวณหาตำแหน่งจุดศูนย์กลางใหม่ของแต่ละกลุ่มจากค่าของข้อมูลที่เป็นสมาชิกในกลุ่มนั้น [5, 6]



ภาพที่ 2: ภาพแสดงอัลกอริทึมของเคมีน

2.5 งานวิจัยที่เกี่ยวข้อง

การค้นหาสินค้าและบริการในท้องถิ่นบนโทรศัพท์เคลื่อนที่ โดยประยุกต์ใช้ระบบสารสนเทศทางภูมิศาสตร์ร่วมกับระบบระบุตำแหน่งบนพื้นโลกและเทคโนโลยีเว็บเซอร์วิส

ประยุกต์ใช้ระบบสารสนเทศทางภูมิศาสตร์ร่วมกับระบบระบุตำแหน่งบนพื้นโลกและเทคโนโลยีเว็บเซอร์วิสโดยประยุกต์ใช้ระบบสารสนเทศทางภูมิศาสตร์ในการแสดงภาพตำแหน่งร้านค้าและตำแหน่งผู้ใช้ในรูปแบบแผนที่ ใช้ระบบระบุตำแหน่งบนพื้นโลกในการระบุตำแหน่งปัจจุบันของผู้ใช้และเทคโนโลยีเว็บเซอร์วิสเป็นตัวกลางในการเชื่อมต่อระหว่างแอปพลิเคชันบนโทรศัพท์เคลื่อนที่กับฐานข้อมูล ด้วยคุณสมบัติของเว็บเซอร์วิส ทำให้ผลลัพธ์ในการค้นหาเป็นปัจจุบันตลอดเวลา [7]

งานบริการลูกค้าสัมพันธ์ได้ทำระบบภูมิสารสนเทศด้วยกูเกิลแมพ เพื่อการบริหารลูกค้าสัมพันธ์ ช่วยสนับสนุนงานด้านการบริหารลูกค้าสัมพันธ์ โดยอิงหลักการแบ่งกลุ่มลูกค้าเชิงภูมิศาสตร์ โดยนำหลักการระบบสารสนเทศภูมิศาสตร์ (GIS) และการบริหารลูกค้าสัมพันธ์(CRM) [8]

ระบบคำแนะนำในการเลือกซื้อเลนส์สำหรับกล้องดีเอสแอลอาร์ โดยการใช้เทคนิค Constance-based Filtering ร่วมกับเทคนิค Collaborative Filtering ทำการแนะนำเลนส์ที่มีทั้งหมด 283 เลนส์ ที่มีความคล้ายคลึงกับคุณลักษณะของเลนส์กล้องที่ผู้ใช้ต้องการ อัลกอริทึมในการหาความคล้ายคลึงคือการทำนายด้วยวิธีการหาสมาชิกที่ใกล้กันที่สุด K-NN (K-Nearest Neighbour) [9]

ระบบแนะนำร้านอาหารด้วยเทคนิคการกรองเชิงร่วมมือและเทคนิคการกรองแบบอิงเนื้อหากรณีศึกษา จังหวัดพระนครศรีอยุธยา ใช้เทคนิคการกรองเชิงร่วมมือหาค่าความคล้ายคลึงด้วยวิธี Pearson Correlation Coefficient และเทคนิคการกรองแบบอิงเนื้อหาด้วยวิธี Euclidean Distance ในการคำนวณหาระยะห่างของข้อมูลบนพื้นฐานของ K-Nearest Neighbour เพื่อแนะนำร้านอาหารได้ตรงตามความต้องการของผู้ใช้รวมถึงแสดงแผนที่ของร้านอาหารด้วย Google Maps API [10]

3. วิธีดำเนินการวิจัย

วิธีดำเนินการวิจัยของระบบแนะนำสถานที่ท่องเที่ยวอุทยานแห่งชาติทางภาคใต้ของประเทศไทยด้วยกูเกิลแมพ เอฟไอ บนมือถือ สมาร์ทโฟน ผู้วิจัยได้ศึกษาข้อมูล และทฤษฎีที่เกี่ยวข้องเพื่อนำมาแก้ปัญหาของสถานที่ท่องเที่ยวที่เป็น

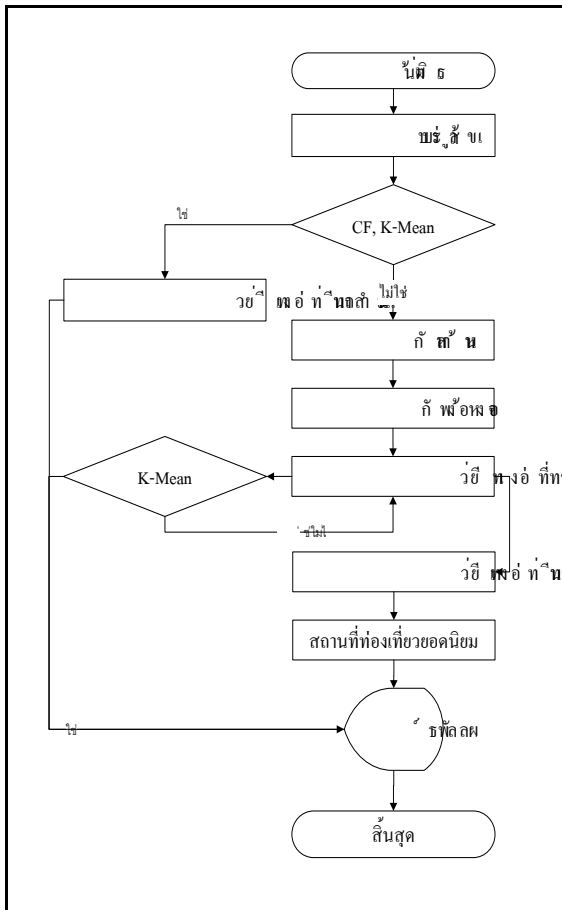
อุทยานแห่งชาติของภาคใต้ในปัจจุบัน ซึ่งผู้วิจัยจัดทำสารนิพนธ์มีวิธีดำเนินการวิจัยดังต่อไปนี้

3.1 ศึกษาและรวบรวมข้อมูล

ได้ทำการศึกษาและรวบรวมข้อมูลของระบบโดยศึกษาจากข้อมูลเดิมสามารถกำหนดระดับความชอบของสถานที่ท่องเที่ยวทางอุทยานเป็นมาตรฐานได้ 5 อันดับ ระดับความชอบของ Rating Scale ตั้งแต่ (1-5) ไม่ชอบมาก, ไม่ชอบ, เฉยๆ, ชอบและชอบมาก

3.2 การวิเคราะห์ระบบ

จากการศึกษาความต้องการและรวบรวมข้อมูล สามารถวิเคราะห์โครงสร้างการทำงานของระบบโดยรวม แสดงลำดับขั้นตอนการทำงานของระบบดัง ภาพที่ 3

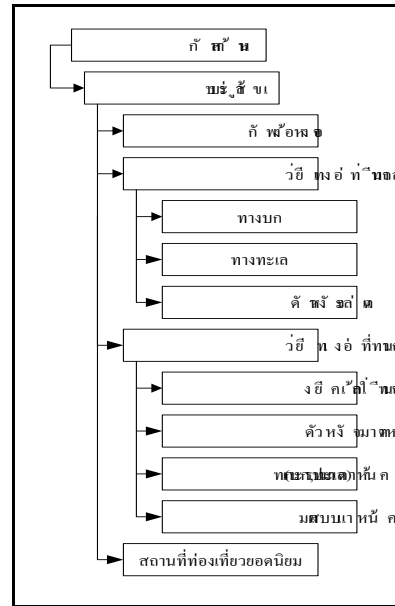


ภาพที่ 3: ขั้นตอนการทำงานของระบบ

3.3 การออกแบบระบบ

การดำเนินงานภายในระบบจะประกอบไปด้วย 6 ส่วนหลักดังภาพที่ 4 คือ ส่วนของสมาชิก ส่วนของรายละเอียดสถานที่

ท่องเที่ยว ส่วนของการจองที่พัก ส่วนของการค้นหาสถานที่ท่องเที่ยวและส่วนของสถานที่ท่องเที่ยวยอดนิยม



ภาพที่ 4: แผนที่เว็บไซต์ระบบแนะนำสถานที่ท่องเที่ยวอุทยานแห่งชาติทางภาคใต้ของประเทศไทยด้วยกูเกิลแมพ เอพีไอ บนมือถือสมาร์ทโฟน

3.4 การพัฒนาระบบ

การพัฒนาระบบแนะนำสถานที่ท่องเที่ยวอุทยานแห่งชาติทางภาคใต้ของประเทศไทย ด้วยกูเกิลแมพ เอพีไอ บนมือถือสมาร์ทโฟน การพัฒนาระบบใช้โปรแกรม Eclipse IDE for Java Developers ในการพัฒนาระบบ ใช้ภาษาจาวาในการควบคุมการทำงานของระบบ พัฒนาระบบปฏิบัติการแอนดรอยด์ เวอร์ชัน 4.2 เป็นต้นไป ใช้กูเกิลแมพ เอพีไอ ในการดึงข้อมูลบนแผนที่เพื่อนำมาช่วยแนะนำสถานที่ท่องเที่ยวและใช้ เฟซบุ๊ก เอพีไอ มาช่วยดึงข้อมูลของสถานที่ท่องเที่ยวเพื่อนำมาช่วยในการจัดเรตติ้งให้กับสถานที่ท่องเที่ยวต่างๆ โดยการนำรายละเอียดการวิเคราะห์และออกแบบระบบที่ได้ออกแบบไว้มาพัฒนาให้ได้ระบบที่สมบูรณ์

3.5 การทดสอบความถูกต้องของระบบ

ระบบแนะนำสถานที่ท่องเที่ยวอุทยานแห่งชาติทางภาคใต้ของประเทศไทย ด้วยกูเกิลแมพ เอพีไอ บนมือถือ สมาร์ทโฟน ทดสอบโดยการทดสอบแบบกล่องดำซึ่งเป็นการทดสอบการ

ทำงานของระบบ โดยรวมทั้งหมดว่ามีกระบวนการทำงาน ถูกต้องตามวัตถุประสงค์หรือไม่สามารถเรียกได้อีกอย่างว่า Functional Testing คือ การทดสอบระบบโดยมองข้ามต่อกลไกภายในของระบบและมุ่งเน้นไปที่ผลลัพธ์ที่ออกมาหลังจากการตอบสนองของระบบอย่างเดียว โดยเลือกข้อมูลนำเข้าและเงื่อนไขการประมวลผลข้อมูลที่นำมาทดสอบเป็นทั้งข้อมูลที่ถูกต้องและไม่ถูกต้อง แล้วทำการกรอกข้อมูลลงไปเสมือนกับการทำงานจริง ๆ เพื่อดูว่าโปรแกรมสามารถทำงานและให้ผลตามที่ผู้ใช้งานต้องการหรือไม่ซึ่งในการทดสอบนี้จะแบ่งการทดสอบโปรแกรมออกเป็น 3 ส่วนหลัก ๆ ดังตารางที่ 1

ตารางที่ 1 : ตารางทดสอบความถูกต้องของระบบ

ลำดับ	ฟังก์ชัน	รายละเอียด	ข้อมูลนำเข้า	ผลลัพธ์
1	ทดสอบการหาห้องพักของอุทยาน			
	จองห้องพัก	ฟังก์ชันการจองห้องพัก	สถานที่	สามารถจองห้องพักได้จริง
2	ทดสอบการค้นหาสถานที่ท่องเที่ยว			
	ค้นหาสถานที่ท่องเที่ยว	ฟังก์ชันการค้นหาสถานที่ท่องเที่ยว	สถานที่	แสดงที่ตั้งสถานที่ท่องเที่ยวได้อย่างถูกต้อง
3	ทดสอบการหาข้อมูลรายละเอียดของสถานที่ท่องเที่ยว			
	ข้อมูลสถานที่ท่องเที่ยว	ฟังก์ชันข้อมูลสถานที่ท่องเที่ยว	สถานที่	ได้รายละเอียดสถานที่ท่องเที่ยวที่ต้องการ

4. ผลการดำเนินงาน

4.1 ผลการพัฒนาาระบบ

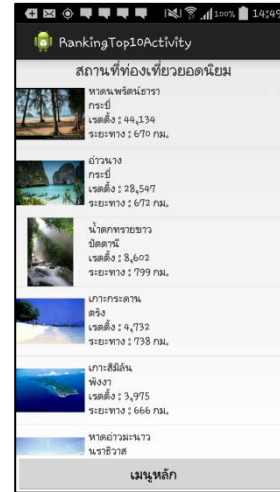
4.1.1 หน้าเมนูหลัก



ภาพที่ 5: หน้าหลักของระบบ

หน้าเมนูหลักจะประกอบไปด้วยเมนูการจองห้องพัก เมนูข้อมูลสถานที่ท่องเที่ยว เมนูค้นหาสถานที่ท่องเที่ยว และสถานที่ท่องเที่ยวยอดนิยม ดังภาพที่ 4

4.1.2 เมนูสถานที่ท่องเที่ยวยอดนิยม



ภาพที่ 6: สถานที่ท่องเที่ยวยอดนิยม

เมนูสถานที่ยอดนิยมจะประกอบด้วยรายชื่อสถานที่ท่องเที่ยวที่คนส่วนใหญ่นิยมไปท่องเที่ยวโดยจะเรียงลำดับจากความนิยมมากที่สุดไปหาน้อย ดังภาพที่ 5

4.2 ผลประเมินความพึงพอใจ

ตารางที่ 2 : ผลการประเมินความพึงพอใจของผู้ใช้งาน

รายการประเมิน	ระดับคุณภาพ	
	\bar{x}	S.D
1.ด้านประสิทธิภาพ(Performance Test)	4.5	0.60
2.ด้านความถูกต้อง(Reliability Test)	4.8	0.64
3.ด้านความสะดวกในการใช้งาน(Usability Test)	4.3	0.61
ค่าเฉลี่ยรวม	4.5	0.02

จากการประเมินความพึงพอใจของระบบมีค่าเฉลี่ยโดยรวมเท่ากับ 4.5 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.02 ดังนั้นระบบมีความพึงพอใจอยู่ในระดับดี

ตารางที่ 3 : ผลการประเมินประสิทธิภาพโดยผู้เชี่ยวชาญ

รายการประเมิน	ระดับคุณภาพ	
	\bar{x}	S.D
1.ด้านประสิทธิภาพ(Performance Test)	4.46	0.51
2.ด้านความถูกต้อง(Reliability Test)	4.46	0.51
3.ด้านความสะดวกในการใช้งาน(Usability Test)	4.41	0.66
4.ด้านความสามารถในหน้าที่การทำงานของระบบ (Function Test)	4.58	0.58
5.ด้านความปลอดภัย(Security Test)	4.55	0.52
ค่าเฉลี่ยรวม	4.49	0.06

จากการประเมินประสิทธิภาพโดยผู้เชี่ยวชาญระบบมีค่าเฉลี่ยโดยรวมเท่ากับ 4.38 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.06 ดังนั้นระบบมีคุณภาพอยู่ในระดับดี

5. สรุปผล

การพัฒนากระบวนการแนะนำสถานที่ท่องเที่ยวอุทยานแห่งชาติทางภาคใต้ของประเทศไทยโดยใช้กูเกิลแมพ เอพีไอและใช้เทคนิคตัวกรองเชิงร่วมมือร่วมกับเคมีน โดยการใช้ทั้ง 2 เทคนิคนี้ควบคู่กันมีข้อดีคือช่วยแนะนำข้อมูลให้ตรงกับความต้องการมากยิ่งขึ้นซึ่งการพัฒนากระบวนการแนะนำนี้มีความน่าสนใจและสะดวกมากขึ้นเป็นการตอบโจทย์ความต้องการของกลุ่มนักท่องเที่ยวโดยตรง และยังช่วยส่งเสริมการท่องเที่ยวของประเทศไทยอีกด้วย ผลจากการสำรวจความพึงพอใจ โดยทำการประเมินความพึงพอใจของผู้ใช้งานต่อระบบที่พัฒนาขึ้นผลที่ได้คือผู้ใช้งานมีความพึงพอใจอยู่ในระดับดีและประเมินประสิทธิภาพโดยผู้เชี่ยวชาญได้ผลออกมาว่าอยู่ในระดับดีจึงสามารถนำระบบแนะนำนี้ไปใช้งานได้อย่างมีประสิทธิภาพ

เอกสารอ้างอิง

- [1] อุทยานแห่งชาติ สัตว์ป่า และพันธุ์พืช กระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อม [ออนไลน์]. Available from <http://www.dnp.go.th/> [2556, กรกฎาคม 3].
- [2] ชยพล น้อยคัด. ระบบแนะนำการเลือกบริโภคอาหารที่มีประโยชน์ต่อสุขภาพโดยใช้วิธีการกรองแบบอ้างอิงตามเนื้อหาร่วมกับการวิเคราะห์จากดัชนีมวลกายและจำนวนพลังงานที่ร่างกายต้องการ ปัญหาพิเศษ สาขาวิชา เทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2555.
- [3] ระบบปฏิบัติการแอนดรอยด์.[ออนไลน์]. Available from: <http://www.th.wikipedia/wiki/แอนดรอยด์> (ระบบปฏิบัติการ) [2556, กรกฎาคม 21].
- [4] กูเกิลแมพ เอพีไอ.[ออนไลน์]. Available from: <http://www.Googlemaps.in.th> [2556, กรกฎาคม 21]

- [5] กฤษณะ ไวยมัย. “คลังข้อมูลและการทำเหมืองข้อมูล” เอกสารคำสอน วิชา 214554. คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์, 2549.
- [6] J. Han and M. Kamber. “Data Mining: Concept and Techniques: 2nd.” USA: Morgan Kaufmann Publication, 2549.
- [7] สุพัฒน์ สุขเกษม. ระบบค้นหาสินค้าและบริการในท้องถิ่นบนโทรศัพท์เคลื่อนที่โดยประยุกต์ใช้ระบบสารสนเทศทางภูมิศาสตร์ร่วมกับระบบระบุตำแหน่งบนพื้นโลกและเทคโนโลยีเว็บเซอร์วิส วิทยาสตรมหาบัณฑิต สาขาวิชาระบบสารสนเทศเพื่อการจัดการ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2552.
- [8] จรัสศรี เจ็งวัฒนพงศ์. ระบบภูมิสารสนเทศด้วย กูเกิลแมพเพื่อการบริหารลูกค้าสัมพันธ์. วิทยาสตรมหาบัณฑิต แผนก วิชา ระบบสารสนเทศเพื่อการจัดการ สาขาวิชาเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2554.
- [9] สุภรัตน์ รัศมีเทศ. ระบบให้คำแนะนำในการเลือกซื้อเลนส์สำหรับกล้องดีเอสแอลอาร์ การประชุมวิชาการ NCCIT สาขาวิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2554.
- [10] บัณฑิต พุทธสร. ระบบแนะนำร้านอาหารด้วยเทคนิคการกรองเชิงร่วมมือและ เทคนิคการกรองแบบอิงเนื้อหา กรณีศึกษา จังหวัดพระนครศรีอยุธยา สาขาวิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2554.

ระบบแนะนำสาขาในการเรียนต่อที่ใช้เรเพอทอรีกริด

A Recommendation System to Choose a Study Program based on Repertory Grid

เดชาวุฒิ วานิชสรรพ (Dechawut Wanichsan)¹ และสาธิต สุวรรณเวช (Sathit Suwannawach)²

คณะวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏรำไพพรรณี

¹kook260g@hotmail.com, ²thequese@hotmail.com

บทคัดย่อ

วัตถุประสงค์หลักของงานวิจัยชิ้นนี้คือเสนอขั้นตอนวิธีการสร้างระบบแนะนำสาขาในการเรียนต่อที่ทำงานบนพื้นฐานจากเทคนิคเรเพอทอรีกริด (Repertory Grid) ซึ่งสามารถช่วยให้อ่านคำแนะนำในการเลือกสาขาสำหรับการเรียนอย่างตรงความต้องการของผู้ใช้อย่างแท้จริงตามความรู้ของผู้ใช้งาน เหตุผลที่ระบบแนะนำสาขาการเรียนที่งานวิจัยชิ้นนี้นำเสนอไม่ได้ใช้งานขั้นตอนวิธียอดนิยมที่งานวิจัยส่วนมากใช้นั้นคือวิธีการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม (Collaborative Filtering) ด้วยแนวคิดที่ว่าสำหรับระบบแนะนำที่เกี่ยวข้องกับการศึกษาแล้วขั้นตอนวิธีที่ใช้ควรสนใจความรู้ของผู้ใช้งานเป็นหลัก มากกว่าที่จะสนใจว่าผู้ใช้งานชอบหรือไม่ชอบสิ่งใด หากผลของการแนะนำไม่มีความแม่นยำแล้วจะส่งผลเสียต่ออนาคตของผู้ใช้งานระบบโดยตรง นอกจากนี้งานวิจัยชิ้นนี้ยังได้เสนอขั้นตอนวิธีในการคำนวณค่าขีดแบ่ง (Threshold) โดยอัตโนมัติเพื่อใช้จำกัดจำนวนคำแนะนำที่เหมาะสมที่สุดให้ผู้ใช้งานอีกด้วย

คำสำคัญ: ระบบแนะนำ เรเพอทอรีกริด วิธีการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม

Abstract

The main objective of this research was to propose an algorithm to develop a recommendation system to choose a study program based on repertory grid technique. The system could provide users accurate guidance about choosing study programs. The reason this proposed algorithm was not based on one of the popular algorithms, that is, a collaborative

filtering method was that an idea which an educational recommendation system should work and rely on an individual user's knowledge rather than focusing preference of any other users who shared similar characteristics to the user. Without accurate guidance from the system, a user would face with some troubles in the future. Moreover, this research also proposed an algorithm to automatically define a threshold value to provide a user with the best recommendation.

Keywords: Recommendation System, Repertory Grid, Collaborative Filtering.

1. บทนำ

ในปัจจุบันคอมพิวเตอร์มีบทบาทที่สำคัญต่อชีวิตประจำวันของมนุษย์ ซอฟต์แวร์ต่าง ๆ ซึ่งเป็นส่วนหนึ่งของระบบคอมพิวเตอร์ได้รับการพัฒนาให้มีความชาญฉลาดมากขึ้นในทุกวันบนพื้นฐานของการพัฒนาขั้นตอนวิธี (Algorithm) ที่มีความชาญฉลาด โดยขั้นตอนวิธีเหล่านี้ทำงานอยู่เบื้องหลังระบบสารสนเทศที่ชาญฉลาด (Intelligent System) ได้รับการพัฒนาขึ้นเพื่อใช้ประโยชน์ในชีวิตประจำวันหลากหลายด้าน เช่น ระบบรู้จำเสียงพูดของมนุษย์โดยอัตโนมัติ (Automatic Speech Recognition System) ระบบระบุตัวตนด้วยชีวมาตร (Biometric Identification and Verification System) และระบบให้คำแนะนำต่าง ๆ กับผู้ใช้งาน (Recommendation System) เป็นต้น ระบบการให้คำแนะนำเป็นระบบที่คอมพิวเตอร์สามารถเสนอข้อมูลให้ผู้ใช้งานจากการคาดเดาส่งผู้ใช้งานเป็นหรือสิ่งที่ผู้ใช้งานกำลังสนใจอยู่เพื่อให้ผู้ใช้งานสามารถตัดสินใจได้ง่ายยิ่งขึ้น ตัวอย่างที่เห็นได้ชัดสำหรับคนที่ใช้งาน

คอมพิวเตอร์เป็นประจำได้แก่ เว็บไซต์กูเกิล (www.google.com) ที่มีข้อความเชื่อมโยงไปยังรายการค้นหาที่มีความเกี่ยวข้อง (Similar Page) มีงานวิจัยหลายชิ้นที่มีหัวข้อเกี่ยวกับระบบแนะนำในด้านต่าง ๆ เช่น คนตรี [1] สินค้าอิเล็กทรอนิกส์ [2] รายการวิดีโอบนยูทูป ภาพยนตร์ [3] การค้นหาบนเครือข่ายทางสังคม (Social Network) และด้านการศึกษา [4][5] เป็นต้น

ระบบให้คำแนะนำทุกระบบที่ใช้งานอยู่ในปัจจุบันจะมีส่วนประกอบหลักที่สำคัญอยู่ที่หมุดสี่ส่วนได้แก่ (1) ส่วนที่เก็บข้อมูลทั่วไปของผู้ใช้งานระบบ ส่วนนี้อาจใช้เป็นข้อมูลเสริมในการประมวลผล เช่น ชื่อ-นามสกุล ประวัติส่วนตัวต่าง ๆ เป็นต้น (2) ส่วนการป้อนข้อมูล ส่วนนี้เป็นส่วนรับข้อมูลเข้า (Input) โดยข้อมูลที่รับมานั้นมีทั้งข้อมูลที่รับมาโดยตรง เช่น การตั้งคำถามโดยต้องการคำตอบที่ชัดเจนเป็นกรณี ๆ ไป การถามความนิยมที่ต้องการคำตอบเป็นระดับคะแนนที่คล้ายกับการถามแบบสอบถาม เป็นต้น และข้อมูลที่ระบบรับเข้ามาทางอ้อมเช่น สังเกตและสกัดข้อมูลมาจากพฤติกรรมของผู้ใช้งาน เช่น ประวัติการเยี่ยมชมข้อมูลเว็บเพจหน้าต่าง ๆ ประวัติการซื้อสินค้า ช่วงเวลาที่ใช้งานระบบ เป็นต้น (3) ส่วนขั้นตอนวิธีหรือสมการที่ใช้คำนวณ ส่วนนี้เป็นส่วนที่เป็นที่สำคัญมากที่สุดเมื่อเทียบกับส่วนอื่น ๆ ทุกส่วนเพราะประสิทธิภาพของคำแนะนำที่ได้รับจากระบบแนะนำจะมีคุณภาพเพียงใดขึ้นอยู่กับการทำงานของส่วนนี้เป็นหลัก โดยขั้นตอนวิธีที่เป็นที่นิยมในปัจจุบันคือวิธีการกรองข้อมูลแบบพึ่งพาผู้ใช้งานร่วม (Collaborative Filtering) [2][3][4][5] และ (4) ส่วนแปลผลและนำเสนอให้ผู้ใช้งาน ส่วนนี้เป็นส่วนแสดงออก (Output) ที่ช่วยนำเสนอผลลัพธ์ให้มีความง่ายต่อการทำความเข้าใจ

ในบริบทเกี่ยวกับการศึกษา การเลือกสาขาหรือวิชาเอกเพื่อศึกษาต่อมีอยู่หลายกรณี เช่น หลังจากระดับมัธยมศึกษาตอนต้นและผู้เรียนต้องเลือกว่าตนเองควรศึกษาต่อสายสามัญหรือควรเบนเข็มไปยังสายอาชีพ หากต้องการเรียนสายสามัญ ควรเรียนต่อในสายวิทยาศาสตร์หรือสายศิลป์ เป็นต้น การเลือกสาขาหรือสายการเรียนที่ไม่เหมาะสมกับตนเองเป็นเหตุให้ผู้เรียนพบอุปสรรคในการเรียน ในกรณีที่เลวร้ายที่สุดคือผู้อาจทำให้ต้องพักอาจเรียนและต้องลาออกจากการเรียน โดยที่ไม่สำเร็จ

การศึกษา การเลือกสาขาเรียนต่อนั้นผู้เรียนส่วนมากยังขาดข้อมูล รวมถึงยังมีประสบการณ์ที่ไม่เพียงพอทำให้ไม่สามารถทราบได้ว่าสาขาวิชาใดมีความเหมาะสมกับความถนัดและความรู้ที่ของตนเองมีอยู่ ผู้เรียนส่วนมากจึงตัดสินใจเลือกสาขาวิชาโดยใช้ความรู้สึก รวมทั้งเพื่อน หรือผู้ปกครองเป็นผู้ชี้แนะหลัก ทำให้เมื่อเข้าไปเรียนแล้ว จึงค้นพบตัวเองในภายหลังว่าไม่เหมาะสมกับสาขาวิชาเหล่านั้น

ด้วยปัญหาดังกล่าวแสดงให้เห็นถึงความจำเป็นในการพัฒนางานวิจัยที่เกี่ยวกับระบบคอมพิวเตอร์ที่สามารถให้คำแนะนำเกี่ยวกับการศึกษาต่อได้ ระบบนี้จะช่วยให้ผู้เรียนทราบถึงความถนัดที่แท้จริงว่าตนเองมีความเหมาะสมกับการเรียนในสาขาใด เพื่อให้ผู้เรียนมีแนวโน้มที่จะประสบความสำเร็จในการเรียนและเรียนได้อย่างมีความสุข งานวิจัยส่วนมากที่ใช้เพื่อสร้างระบบแนะนำ จะพัฒนาบนพื้นฐานของวิธีการกรองข้อมูลแบบพึ่งพาผู้ใช้งานร่วม (Collaborative Filtering) หลักการทำงานโดยสรุปของวิธีนี้คือขั้นตอนวิธีจะทำการเลือกผู้ใช้งานในระบบคนอื่น ๆ ที่มีลักษณะหรือมีพฤติกรรมใกล้เคียงกับผู้ใช้งานมากที่สุด และจะใช้ข้อมูลเหล่านั้นเพื่อทำนายหรือให้คำแนะนำกับผู้ใช้งาน

อย่างไรก็ตามแม้ว่าวิธีการกรองข้อมูลแบบพึ่งพาผู้ใช้งานจะเป็นที่นิยมและสามารถใช้งานเพื่อสร้างระบบแนะนำได้หลากหลายบริบท เช่น ใช้งานบนระบบซื้อขายสินค้าออนไลน์ และระบบแนะนำภาพยนตร์เรื่องที่ชอบ เป็นต้น แต่สำหรับระบบแนะนำที่เกี่ยวข้องกับการศึกษาแล้ว คำแนะนำที่ระบบสร้างขึ้นควรพัฒนาจากความรู้หรือความถนัดของผู้ใช้ มากกว่าพัฒนาจากจากความชอบของผู้ใช้ โดยเฉพาะอย่างยิ่ง วิธีการกรองข้อมูลแบบพึ่งพาผู้ใช้งานที่มีหลักการทำงาน โดยใช้ความชอบของผู้ใช้คนอื่น ๆ ที่มีลักษณะคล้ายกับผู้ใช้งานหลักและทำการอนุมานค่าความชอบนั้นจนได้ผลลัพธ์เป็นคำแนะนำจากระบบ

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ระบบแนะนำที่สร้างโดยใช้วิธีนี้เป็นหลักสามารถให้คำแนะนำเกี่ยวกับข้อมูลให้กับผู้ใช้งานได้ ผู้ใช้ที่อยู่ในระบบจะทำการประเมินความชอบหรือไม่ชอบต่อชิ้นข้อมูล ซึ่งข้อมูลเหล่านี้จะบอกเป็นนัยถึงข้อมูลส่วนตัวของผู้ใช้ จากนั้นวิธีการนี้

จะใช้ข้อมูลทางสถิติในการค้นหาข้อมูลผู้ใช้งานคนอื่น ๆ ที่มีลักษณะบางอย่างที่มีความชอบคล้ายกับผู้ใช้หลัก ซึ่งข้อมูลเหล่านี้จะได้รับการวิเคราะห์ผล เพื่อสร้างคำแนะนำให้ตรงตามความต้องการของผู้ใช้งานหลักต่อไป โดยปกติวิธีการที่ใช้ในการวัดค่าความคล้ายคลึงของข้อมูล (Similarity Measure) ของผู้ใช้งานแต่ละคนที่นิยมใช้กันคือ สหสัมพันธ์ของเพียร์สัน (Pearson Correlation) การหาความคล้ายเชิงมุม (Cosine Similarity) และการวัดระยะส่วนกลับของยูคลิด (Inverse Euclidean Distance) แต่เนื่องจากพื้นที่เขียนมีจำกัด จึงไม่ขอให้รายละเอียดสมการของวิธีการทั้งสาม อ่านรายละเอียดเพิ่มเติมได้ที่ [2][3] จากนั้นจึงทำการเลือกผู้ใช้งานที่ใกล้เคียงกับผู้ใช้เป้าหมายมากที่สุด ซึ่งอาจมีหนึ่งรายการหรือหลายรายการก็ได้ วิธีการที่ง่ายที่สุดคือขั้นตอนวิธีหาเพื่อนบ้านที่ใกล้ที่สุด k ตัว (K-nearest Neighbor) จากนั้นจึงหาค่าเฉลี่ยของค่าความชอบที่ได้จากเพื่อนบ้านเหล่านั้นเพื่อสร้างคำแนะนำให้กับผู้ใช้หลัก อย่างไรก็ตาม ข้อจำกัดหลายอย่างที่ทำให้ประสิทธิภาพของวิธีนี้ลดน้อยลงได้แก่ (1) เมื่อผู้ใช้งานที่อยู่ในระบบมีจำนวนไม่เพียงพอทำให้ไม่สามารถทำการจับคู่ความคล้ายคลึงกับผู้ใช้งานอื่น ๆ ได้ (2) แม้ระบบจะมีผู้ใช้งานจำนวนมาก แต่ถ้าผู้ใช้ส่วนใหญ่กำหนดค่าความชอบให้สินค้าเป็นบางรายการ (ทำให้ข้อมูลมีสภาพเบาบาง) ก็ทำให้เป็นเรื่องยากที่หาผู้ใช้งานที่มีความคล้ายอย่างแท้จริงกับผู้ใช้หลัก นอกจากนี้ยังทำให้เวลาที่ใช้ในการประมวลผลนานยิ่งขึ้นด้วย และ (3) ซึ่งเป็นข้อจำกัดที่สำคัญที่สุดคือ ระบบแนะนำที่เกี่ยวข้องกับการศึกษา [4][5] ควรใช้ความรู้ของผู้ใช้งานเป็นหลัก มากกว่าใช้เก็บข้อมูลผู้ใช้ว่าชอบหรือไม่ชอบ หรือนำข้อมูลของผู้ใช้คนอื่นมาอนุมานคำตอบให้ผู้ใช้หลัก

3. วิธีดำเนินงานวิจัย

จากปัญหาที่กล่าวไปแล้วในหัวข้อที่ผ่านมา งานวิจัยชิ้นนี้จึงมุ่งที่จะเสนอวิธีการพัฒนาระบบแนะนำสาขาในการเรียนต่อที่ ให้คำแนะนำผู้ใช้จากความถนัดและความรู้ที่ผู้ใช้งานมีอยู่ วิธีดำเนินการวิจัยของงานชิ้นนี้แบ่งเนื้อหาออกเป็นสามส่วนหลัก ส่วนแรกคือโครงสร้างของระบบแนะนำสาขาในการเรียนต่อโดยใช้เทคนิคเรพเพทอริกิด ที่ช่วยให้เห็นบริบทของระบบว่ามีความเกี่ยวข้องกับสิ่งใดและมีองค์ประกอบอะไรบ้าง ส่วนที่

สองคือรายละเอียดของเรพเพทอริกิด และส่วนสุดท้ายคือขั้นตอนวิธีและสมการที่ใช้ในการคำนวณค่าต่าง ๆ เพื่อให้ระบบสร้างคำแนะนำให้กับผู้ใช้งาน

3.1 โครงสร้างของระบบแนะนำสาขาในการเรียนต่อโดยใช้เทคนิคเรพเพทอริกิด

โครงสร้างของระบบแสดงได้ดังภาพที่ 1 ที่มีผู้ใช้งานหลักอยู่สองประเภทได้แก่ ผู้เชี่ยวชาญหรือครู และผู้ใช้ระบบหรือนักเรียน ในที่นี้ผู้เชี่ยวชาญอาจมีคนเดียวหรืออาจมีหลายคนก็ได้ หน้าที่หลักของผู้เชี่ยวชาญคือการวิเคราะห์ข้อสอบด้วยการให้ค่าน้ำหนักบนเรพเพทอริกิด (แสดงในหัวข้อที่ 3.2) หน้าที่ของผู้ใช้ระบบคือทำข้อสอบบนระบบและรับคำแนะนำที่ได้จากระบบ โดยผลการทำข้อสอบจะได้รับการบันทึกไว้ในฐานข้อมูลเพื่อใช้เป็นข้อมูลนำเข้าร่วมกับค่าน้ำหนักที่ได้จากผู้เชี่ยวชาญ สำหรับการคำนวณค่าเพื่อให้คำแนะนำต่อผู้ใช้งาน (สมการที่ใช้คำนวณแสดงในหัวข้อที่ 3.3) นอกจากนี้ยังมีส่วนประกอบอื่น ๆ ที่ยังไม่ได้กล่าวถึงอยู่อีกสองส่วนได้แก่

ข้อสอบซึ่งหมายความว่าข้อสอบที่มีเนื้อหาครอบคลุมทุกสาขาวิชาที่ระบบสามารถให้คำแนะนำกับผู้ใช้หลัก โดยข้อสอบนี้ ผู้เชี่ยวชาญไม่จำเป็นต้องออกข้อสอบและจำนวนข้อในข้อสอบจะมีก็ข้อก็ได้ตรงกับที่มีเนื้อหาครอบคลุมองค์ความรู้ที่จำเป็นของทุกสาขาวิชา เพื่อให้ง่ายต่อการออกแบบข้อสอบควรมีลักษณะเป็นข้อสอบแบบหลายตัวเลือก (Multiple Choices) เพื่อที่จะตัดสินใจได้อย่างชัดเจนว่าคำตอบที่ได้รับจากผู้ใช้ระบบนั้นถูกต้องหรือไม่

อีกส่วนประกอบหนึ่งคือหน้าจอร์บบแบบทดสอบบนคอมพิวเตอร์ ระบบนี้อาจทำงานได้แบบออนไลน์หรือไม่ออนไลน์ก็ได้ขึ้นอยู่กับผู้พัฒนาภาษาคอมพิวเตอร์ที่ใช้ในการพัฒนาโปรแกรม ผู้ใช้งานจะทำข้อสอบผ่านหน้าจอของระบบนี้และรอรับคำแนะนำที่ได้จากระบบ

3.2 องค์ประกอบของเรพเพทอริกิด

เรพเพทอริกิดคือ เทคนิคการสัมภาษณ์ที่ช่วยให้บุคคลสามารถตีความหรือให้ความหมายของความรู้ของตนเองจากประสบการณ์ที่ตนมี เทคนิคนี้เสนอขึ้นครั้งแรกในปี 1955 โดย Gorge Kelly [6] และรายละเอียดการใช้งานเรพเพทอริกิดในแง่มุมอื่น ๆ สามารถอ่านเพิ่มเติมได้จากงานวิจัยเหล่านี้ [7][8]

การที่ระบบแนะนำสาขาการเรียนสามารถให้คำแนะนำได้นั้น ระบบจำเป็นต้องได้รับความรู้จากบุคคลกลุ่มหนึ่งซึ่งมีความสำคัญมากนั่นคือผู้เชี่ยวชาญ ในที่นี้ผู้เชี่ยวชาญอาจเป็นครูหรืออาจารย์ที่มีความรู้ในข้อสอบที่จัดทำขึ้นด้วยว่าสาขาวิชาที่ตนเองมีความเชี่ยวชาญใช้ความรู้ในข้อสอบแต่ละข้อเป็นความรู้พื้นฐานหรือไม่ ตัวอย่างของโครงสร้างของเรเพอทอริกิดใช้ในระบบแนะนำสาขาการเรียนแสดงในตารางที่ 1

ตารางที่ 1: เรเพอทอริกิดที่แสดงความสัมพันธ์ระหว่างข้อสอบแต่ละข้อและสาขาวิชาแต่ละสาขา

	สาขาวิชา m_i					
	m_1	m_2	m_3	...	m_p	
ข้อสอบ ข้อที่ q_j	q_1	$w_{1,1}$	$w_{1,2}$	$w_{1,3}$...	$w_{1,p}$
	q_2	$w_{2,1}$	$w_{2,2}$	$w_{2,3}$...	$w_{2,p}$
	q_3	$w_{3,1}$	$w_{3,2}$	$w_{3,3}$...	$w_{3,p}$
	q_4	$w_{4,1}$	$w_{4,2}$	$w_{4,3}$...	$w_{4,p}$

q_n	$w_{n,1}$	$w_{n,2}$	$w_{n,3}$...	$w_{n,p}$	

จากตารางที่ 1 ตัวแปร q_j หมายถึง ข้อสอบข้อที่ j โดยที่ j มีค่าตั้งแต่ 1, 2, 3, ..., n และตัวแปร m_i หมายถึง สาขาวิชาสาขาที่ i โดยที่ i มีค่าตั้งแต่ 1, 2, 3, ..., p นอกจากนี้ค่าที่แสดงในเซลล์แต่ละเซลล์คือค่าน้ำหนักที่ได้จากผู้เชี่ยวชาญ เช่น $w_{2,4}$ หมายถึงค่าน้ำหนักที่แสดงความสัมพันธ์ระหว่างข้อสอบข้อที่สองกับสาขาวิชาที่สี่ เป็นต้น ค่าน้ำหนักเป็นค่าที่ได้รับจากผู้เชี่ยวชาญโดยช่วงที่เป็นไปได้ในการกรอกค่าน้ำหนักอยู่ระหว่างศูนย์ถึงห้า และความหมายของค่าน้ำหนักแต่ละค่าได้ถูกอธิบายไว้ในตารางที่ 2

ตารางที่ 2: ความหมายของค่าน้ำหนักในเรเพอทอริกิด

ค่าน้ำหนัก	ความหมาย
5	มีความสัมพันธ์ระดับสูงมาก
4	มีความสัมพันธ์ระดับสูง
3	มีความสัมพันธ์ระดับปานกลาง
2	มีความสัมพันธ์ระดับต่ำ
1	มีความสัมพันธ์ระดับต่ำมาก
0	ไม่มีความสัมพันธ์

ตัวอย่างการกำหนดค่าน้ำหนักในเรเพอทอริกิดเป็นดังนี้ สมมุติว่าคณะวิทยาศาสตร์แห่งหนึ่งประกอบด้วยสี่สาขาที่เปิดสอนคือ ฟิสิกส์ เคมี วิทยาศาสตร์สิ่งแวดล้อม เทคโนโลยีสารสนเทศ โดยคณะแห่งนี้ต้องการสร้างระบบแนะนำสาขาเรียนต่อจึงได้สร้างข้อสอบขึ้น และข้อสอบข้อหนึ่งที่มีเนื้อหาเกี่ยวข้องกับวิชาแคลคูลัส ผู้เชี่ยวชาญได้กำหนดค่าน้ำหนักให้ข้อสอบนี้ดังนี้ (5,3,1,1) ซึ่งสามารถอธิบายได้ว่า แคลคูลัสมีความเกี่ยวข้องกับสาขาฟิสิกส์ในระดับสูงมาก มีความเกี่ยวข้องกับสาขาเคมีในระดับปานกลาง และมีความเกี่ยวข้องกับสาขาวิทยาศาสตร์สิ่งแวดล้อม และสาขาเทคโนโลยีสารสนเทศในระดับต่ำมาก

3.3 ขั้นตอนวิธีและสมการต่างๆ ที่ใช้คำนวณค่า

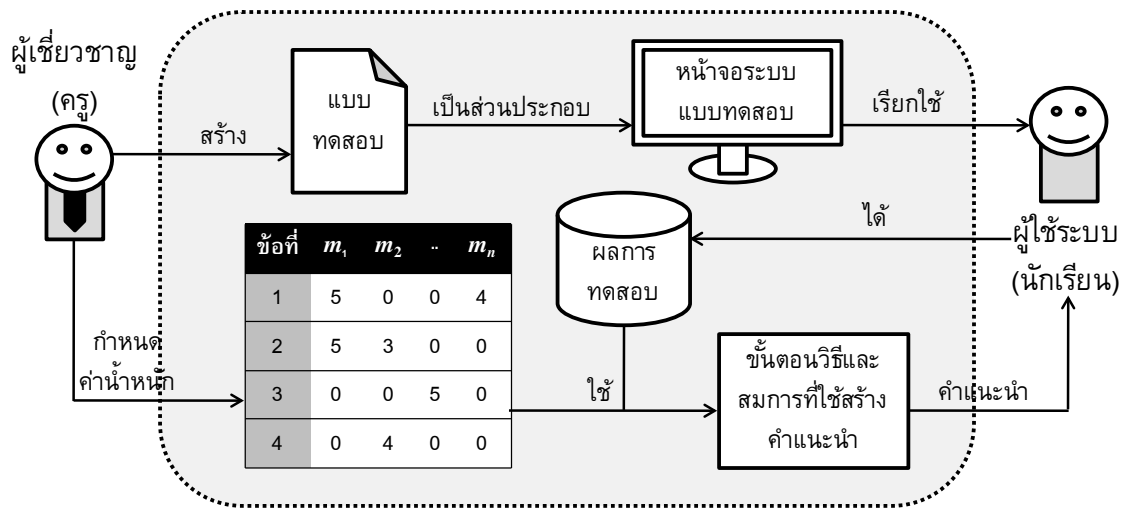
การพิจารณาว่าสาขาในการเรียนใดเป็นสาขาวิชาที่ผู้ใช้ระบบมีความถนัดมากที่สุด สามารถคำนวณค่าความถนัดได้จากสมการที่ 1

$$exp(i) = \left(\sum_{j=1}^n ca_{i,j} \right) / n \tag{1}$$

โดยที่

- i คือสาขาในการเรียนมีค่าตั้งแต่ 1,2,3,... m
- m คือจำนวนสาขาทั้งหมดในระบบแนะนำสามารถรองรับได้
- exp คืออาร์เรย์ที่สมาชิกแต่ละตัวแสดงระดับความถนัดของผู้เรียนต่อสาขาวิชาที่ i
- j คือข้อสอบข้อที่ j จาก n ข้อ
- $ca_{i,j}$ คือค่าน้ำหนักที่ผู้เชี่ยวชาญกำหนดไว้ในเรเพอทอริกิด เฉพาะข้อสอบข้อที่ผู้ใช้งานตอบถูกเท่านั้น

หากค่า $exp(i)$ มีค่ามากกว่าหรือเท่ากับ (\geq) ค่าขีดแบ่งที่กำหนดไว้ (Threshold) นั้นแปลว่า ผู้ใช้งานระบบสามารถเรียนรู้ได้ดีในสาขาวิชาดังกล่าว และสาขาวิชานั้นเป็นสาขาวิชาที่มีความเหมาะสมในการเลือกศึกษาต่อ อย่างไรก็ตามการผู้ควบคุมระบบต้องกำหนดค่าขีดแบ่งด้วยตนเองนับว่าเป็นปัญหาที่ตัดสินใจได้ยากกว่าค่าใดเป็นค่าที่มีความเหมาะสมที่สุดงานวิจัยนี้จึงได้เสนอขั้นตอนวิธีในการประมวลค่าขีดแบ่งโดยอัตโนมัติเพื่อให้ผู้ดูแลระบบแนะนำสาขาการเรียนสามารถใช้งานระบบได้ง่ายยิ่งขึ้น โดยรายละเอียดของขั้นตอนที่เสนอแสดงในภาพที่ 3



ภาพที่ 1: โครงสร้างของระบบแนะนำสาขาในการเรียนต่อที่ใช้เรเพอทริกคิด

ขั้นตอนวิธีนี้ต้องการพารามิเตอร์นำเข้าหนึ่งตัวคือ $nMajor$ ซึ่งคือจำนวนสาขาวิชาที่ผู้ใช้งานถนัดที่สุด เช่น หากกำหนดค่า $nMajor$ เท่ากับสาม หมายความว่าระบบจะเลือกสาขาวิชาที่ผู้ใช้งานถนัดที่สุดจำนวนสามลำดับและแสดงผลให้ผู้ใช้งาน เป็นต้น โดยจำนวนสาขาวิชาที่ผู้ใช้งานถนัดที่สุดที่ระบบสามารถค้นคืนได้จะมีค่ามากกว่าหรือเท่ากับ $nMajor$ ที่กำหนดเป็นข้อมูลนำเข้าของขั้นตอนวิธีเสมอ

ภาพที่ 3: ขั้นตอนวิธีคำนวณค่าขีดแบ่งโดยอัตโนมัติ

```

FUNCTION findThreshold( $nMajor$ )
FOR  $a=1$  TO  $nMajor$ 
     $max = 0$ 
     $pos = 0$ 
    FOR  $i$  IN  $exp$ 
        IF  $exp(i) > max$  THEN
             $max = exp$ 
             $pos = i$ 
        END
    END
     $threshold = exp(i)$ 
    remove  $exp(i)$  from  $exp$ 
END
    
```

4. ตัวอย่างผลการดำเนินงาน

เพื่อให้ผู้อ่านสามารถทำความเข้าใจการสร้างระบบแนะนำสาขาการเรียนต่อที่ใช้เรเพอทริกคิดได้โดยง่าย การยกตัวอย่างการคำนวณค่าจึงเป็นสิ่งที่จำเป็นอย่างยิ่ง ตัวอย่างนี้แสดงให้เห็นว่าระบบแนะนำฯ ซึ่งสามารถรองรับการเลือกสาขาได้ทั้งหมดห้าสาขา แบบทดสอบที่จะใช้วัดความสามารถของผู้ใช้งานเป็น

ข้อสอบจำนวนสิบข้อ นอกจากนี้ค่าน้ำหนักที่ผู้เชี่ยวชาญกำหนดไว้ในเรเพอทริกคิดแสดงดังตารางที่ 4

ตารางที่ 4: เรเพอทริกคิดที่แสดงความสัมพันธ์ระหว่างข้อสอบแต่ละข้อและสาขาวิชาแต่ละสาขา

	สาขาวิชา m_i				
	m_1	m_2	m_3	m_4	m_5
q_1	5	0	0	0	0
q_2	1	0	0	3	5
q_3	1	0	5	0	3
q_4	3	3	3	3	5
q_5	0	0	0	5	0
q_6	0	5	0	0	0
q_7	5	0	0	0	0
q_8	0	4	0	0	4
q_9	0	0	5	0	4
q_{10}	2	2	2	2	5

สมมติว่าผู้ใช้งานระบบคนหนึ่งได้ทำแบบทดสอบจากระบบนี้เสร็จแล้ว และข้อสอบข้อที่ผู้ใช้งานคนนี้ได้ตอบถูกประกอบด้วยข้อสอบข้อที่ 1 2 3 4 7 และ 10 และค่า exp ซึ่งแสดงถึงความถนัดของทุกสาขาวิชาที่คำนวณได้แสดงในตารางที่ 5 เมื่อคำนวณค่าอาร์เรย์ exp แล้ว และกำหนดค่าให้ $nMajor$ มีค่าเท่ากับสาม (ระบบจะเลือกสาขาการเรียนที่เหมาะสมกับ

ผู้เรียนที่สุดเป็นจำนวนสามสาขาวิชาเป็นอย่างน้อย) เพื่อเป็นข้อมูลนำเข้าสู่ของขั้นตอนวิธี *findThreshold* ที่แสดงในตารางที่ 3 ทำให้ค่าขีดแบ่ง (Threshold) ที่คำนวณได้มีค่าเท่ากับ 0.67 นั้นหมายความว่าสาขาวิชาที่ระบบแนะนำสาขาการเรียนจะแนะนำให้กับผู้เรียนคือ สาขาวิชาที่ m_1 , m_3 และ m_5 ตามลำดับ เพราะสาขาวิชาเหล่านี้มีค่า *exp* ที่มากกว่าหรือเท่ากับค่าขีดแบ่ง ซึ่งเท่ากับ 0.67

ตารางที่ 5: ค่าความถนัดที่คำนวณได้

<i>exp</i> (1)	<i>exp</i> (2)	<i>exp</i> (3)	<i>exp</i> (4)	<i>exp</i> (5)
17/17 =	5/14 =	10/15 =	8/13 =	18/26 =
1.00	0.36	0.67	0.62	0.69

5. วิจัยและสรุปผลงานวิจัย

งานวิจัยชิ้นนี้ได้เสนอวิธีการสร้างระบบแนะนำสาขาเรียนต่อที่ใช้เรพอร์ทกริด ซึ่งมีข้อดีมากกว่าระบบแนะนำสาขาการเรียนที่ใช้ขั้นตอนวิธีที่เป็นที่นิยมอยู่ในปัจจุบัน นั่นคือวิธีการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม (Collaborative Filtering) อยู่สามประการ ดังนี้ (1) ระบบนี้จะให้คำแนะนำโดยจะทำการเลือกสาขาการเรียนให้ผู้ใช้จากความรู้ที่ผู้ใช้งานมี ไม่ใช่จากสาขาที่ผู้ใช้งานชอบ ซึ่งนับว่าคำแนะนำที่ได้มีความสมเหตุสมผลมากกว่าระบบที่ใช้วิธีการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม นอกจากนี้ แม้ว่าผู้ใช้งานระบบจะเรียนไม่เก่งนัก ระบบก็ยังสามารถให้คำแนะนำที่เหมาะสมที่สุดกับผู้ใช้เรียนคนนั้น เพราะค่าขีดแบ่งนั้นคำนวณจากค่า *exp* ของสาขาวิชาที่ผู้ใช้คนนั้นถนัดที่สุดแล้ว (2) หากผู้ใช้ระบบคนอื่น ๆ มีจำนวนน้อย ก็จะไม่ส่งผลกระทบใด ๆ ต่อระบบที่สร้างด้วยวิธีการที่เสนอขึ้น เพราะระบบนี้สามารถทำงานได้โดยไม่สนใจว่ามีผู้เคยใช้งานระบบจำนวนกี่คน แตกต่างจากระบบที่ใช้วิธีการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม ที่ประสิทธิภาพของคำแนะนำขึ้นอยู่กับจำนวนผู้ใช้งานคนอื่นเป็นหลัก หากผู้ใช้งานคนอื่น ๆ น้อยคำแนะนำที่ได้ก็มีประสิทธิภาพน้อยเช่นกัน และ (3) หากผู้ใช้งานระบบคนอื่น ๆ มีจำนวนมาก ระบบที่ทำงานบนพื้นฐานวิธีการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วมจะต้องคำนวณค่าความคล้าย (Similarity) ของผู้ใช้งานหลักกับผู้ใช้งานคนอื่น ๆ

ในระบบทุกคน ทำให้ทำงานช้ากว่าระบบแนะนำที่งานวิจัยชิ้นนี้เสนอขึ้น

เอกสารอ้างอิง

- [1] H.C. Chen, and A.L.P. Chen, "A music recommendation system based on music data grouping and user interests", *The ACM International Conference on Information and Knowledge Management*, Atlanta, GA, November 2001.
- [2] วงกต ศรีอุไร และคณะ, "การแทนค่าข้อมูลที่ขาดหายเพื่อแก้ไขปัญหาความเบาบางของข้อมูลในการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม," *วารสารพระจอมเกล้าลาดกระบัง* ปีที่ 16 ฉบับที่ 1 เมษายน พ.ศ.2551 หน้า 44-52.
- [3] U. Suksawatchon, and J. Suksawatchon, "Movie Recommender System Using Collaborative Filtering and Contextual Information", *The 2nd Conference on Knowledge and Smart Technologies (KST 2010)*, Chonburi, Thailand, July 2010.
- [4] D. Lemire, H. Boley, S. McGrath, and M. Ball, "Collaborative filtering and inference rules for context-aware learning object recommendation," *J. Interact. Technol. Smart Educ.* vol. 2, no. 5, pp.179–188, 2005.
- [5] J. Bobadilla, F. Serradilla, and A. Hernando, "Collaborative filtering adapted to recommender systems of e-learning," *Knowledge-Based Systems*, vol. 22, pp. 261–265, 2009.
- [6] F. Fransella, and D. Bannister (1977). *A manual for repertory grid technique*. Academic Press:London, 1977.
- [7] M.G. Hunter, and J.E. Beck, "Using repertory grids to conduct cross-cultural information systems research," *Information Systems Research*, vol. 11, no. 1, pp. 93–101, 2000.
- [8] F.B. Tan, and M.G. Hunter, "The repertory grid technique: a method for the study of cognition in information systems," *MIS Quarterly*, vol. 26, no. 1, pp. 39–57, 2002.

การตรวจจับจุดบกพร่องบนสิ่งทอโดยการหาเส้นขอบภาพร่วมกับซัพพอร์ตเวกเตอร์แมชชีน

Defect Detection in Textile Fabrics with Edge Detection and Support Vector Machines

ทศพล ปราชญ์ปรีชา (Thodsaphon Prachpreecha)¹ และสรกฤษ ศรีเกษม (Sorakis Srikasem)²

¹สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยวงษ์ชวลิตกุล

²สาขาวิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ โรงเรียนนายเรืออากาศ

¹Thodsaphon_pra@vu.ac.th, ²Sorakis_s@rtaf.mi.th

บทคัดย่อ

บทความนี้นำเสนอการตรวจจับจุดบกพร่องอัตโนมัติของผ้าทอ โดยการหาเส้นขอบภาพเพื่อดึงคุณลักษณะเด่นที่แตกต่างของผ้าทอ การลดข้อมูลจะใช้วิธีการวิเคราะห์องค์ประกอบหลักสองมิติ (2DPCA) เพื่อนำไปจำแนกโดยใช้ซัพพอร์ตเวกเตอร์แมชชีน (SVM) ในการประเมินสมรรถภาพของรูปแบบที่ใช้ในการประมวลผล จะใช้รูปภาพผ้าทอจำนวน 120 ภาพเพื่อหาประสิทธิภาพและประสิทธิผล จากผลการทดลอง พบว่า วิธีการดังกล่าว สามารถตรวจจับจุดบกพร่องบนสิ่งทอที่มีความแม่นยำสูงถึง 98.33% จากผลลัพธ์ที่ได้แสดงให้เห็นว่า รูปแบบที่นำเสนอเป็นวิธีที่มีประสิทธิภาพในการตรวจจับจุดบกพร่องบนสิ่งทอ

คำสำคัญ: การตรวจจับจุดบกพร่องอัตโนมัติ การหาเส้นขอบภาพ การวิเคราะห์องค์ประกอบหลักสองมิติ ซัพพอร์ตเวกเตอร์แมชชีน สิ่งทอ

Abstract

This paper proposes an automated defect detection of textile fabrics on classification to classify the deflection/non-deflection fabric. Edge detection is deployed as the major technique for feature extraction of the textile fabric. Data reduction is done by Two-Dimensional Principal Component Analysis (2DPCA), and classification is done by Support Vector Machines. The performance of the scheme is evaluated by using 120 fabric images for efficiency and effectiveness The test

result shows accurate defect detection with effective and robust. The proposed system can defect detection fabrics on the accuracy of 98.33%. The experimental results obtained clearly demonstrate that the proposed scheme is indeed an effective and efficient means for defect detection in textile fabrics.

Keyword: Automated Defect Detection, Edge Detection, 2DPCA, Support Vector Machines, Textile Fabric.

1. บทนำ

ในอุตสาหกรรมสิ่งทอมีความจำเป็นอย่างยิ่งในการตรวจสอบจุดบกพร่องบนพื้นผิว เพื่อควบคุมคุณภาพในการผลิต ในอดีตการตรวจสอบหาจุดบกพร่องบนพื้นผิวผ้าใช้วิธีการแบบง่าย คือ การตรวจสอบด้วยตาเปล่า ซึ่งจะมีความถูกต้องแม่นยำไม่เกิน 80% [1] ดังนั้น จึงมีการวิจัยและพัฒนาอัลกอริทึมสำหรับกระบวนการตรวจสอบจุดบกพร่องมาอย่างต่อเนื่อง มีการคิดค้นเทคนิคและวิธีการใหม่ๆ เพื่อให้เกิดประสิทธิภาพและความแม่นยำอยู่เสมอ โดยอาศัยหลักการประมวลผลข้อมูลภาพแทน [2], [3], [4], [18], [20] ซึ่งมีความถูกต้องแม่นยำสูง เนื่องจากใช้คอมพิวเตอร์แทนการตัดสินใจด้วยมนุษย์

ปัจจุบันมีงานวิจัยที่เกี่ยวข้องในการตรวจหาจุดบกพร่องของพื้นผิวแบ่งออกได้หลายกลุ่ม [5], [10], [11], [15], [17] โดยทั่วไปมักเกี่ยวข้องกับการวิเคราะห์พื้นผิวหรือเนื้อหาของภาพ (Texture Analysis) เป็นการคำนวณหาคุณลักษณะ

ของพื้นผิว (Texture Feature) ที่โดดเด่นมีเอกลักษณ์เฉพาะตัวในการแสดงความเป็นพื้นผิวนั้นๆ ซึ่งจะแตกต่างจากคุณลักษณะพื้นผิวที่ได้จากภาพพื้นผิวอื่นๆ [19] คุณลักษณะของพื้นผิวดังกล่าวสามารถคำนวณได้โดยใช้ความสัมพันธ์ในปริภูมิภายใต้เกณฑ์การกระจายสัญญาณสีเทา การตรวจจับเส้นขอบของภาพ (Edge detection) [17], [20] นอกจากนี้ ซัพพอร์ตเวกเตอร์แมชชีน (SVM) เป็นอีกหนึ่งเทคนิคที่น่าสนใจในการจำแนกพื้นผิว เนื่องจากสามารถดำเนินการง่ายและให้ผลการจำแนกที่มีประสิทธิภาพ งานวิจัยนี้นำเสนอการประยุกต์ใช้การหาเส้นขอบภาพ (Edge detection) ในการดึงคุณลักษณะ ส่วนการลดข้อมูลใช้การวิเคราะห์องค์ประกอบหลักสองมิติ (2DPCA) และซัพพอร์ตเวกเตอร์แมชชีน (SVM) สำหรับการตรวจจับจุดบกพร่องบนสิ่งทอ โดยทำการวิเคราะห์ผ้าทอพื้นเมืองทางด้านกายภาพหรือลักษณะทั่วไปของผ้า ในเรื่องของจุดบกพร่องของผ้าด้วยเทคนิคดังกล่าว เพื่อให้สามารถตรวจจับจุดบกพร่องบนสิ่งทอได้อย่างมีประสิทธิภาพและครอบคลุมปัจจัยที่เกี่ยวข้องให้มากที่สุด เพื่อเพิ่มประสิทธิภาพการผลิตในอุตสาหกรรม ปรับปรุงคุณภาพ ลดเวลา และต้นทุนในการดำเนินงานโดยใช้เทคโนโลยีที่ทันสมัย

2. วรรณกรรมที่เกี่ยวข้อง

การประมวลผลภาพ (Image Processing) มีจุดประสงค์เพื่อพัฒนาและปรับปรุงข้อมูลที่เป็นภาพ เพื่อให้สามารถแปลความหมายได้ดีขึ้น และเพื่อเตรียมข้อมูลฉาก (Scene Data) สำหรับให้คอมพิวเตอร์เข้าใจได้ กระบวนการประมวลผลภาพแบ่งออกได้เป็น 3 กระบวนการ คือ กระบวนการขั้นต้น เป็นการได้มาซึ่งภาพและการปรับปรุงภาพให้ดีขึ้น ได้แก่ การแปลงสัญญาณภาพ การเพิ่มความคมชัดของภาพ การกู้คืนสัญญาณของภาพ เป็นต้น กระบวนการขั้นกลาง คือ ส่วนของการแยกองค์ประกอบและการดึงเอาคุณลักษณะที่สำคัญออกมา และกระบวนการขั้นสูง คือ การตีความหมายและการจดจำภาพ ขั้นตอนที่สำคัญขั้นตอนหนึ่งของการประมวลผลภาพ (Image Processing) คือ การแยกส่วนต่างๆ ของรูปภาพออกจากกันตามลักษณะสำคัญที่เราพิจารณา เพื่อลดจำนวนข้อมูลในรูปภาพ

ที่ไม่จำเป็นในการวิเคราะห์ จัดระเบียบข้อมูลในรูปภาพให้เป็นกลุ่ม และแสดงข้อมูลในรูปที่เข้าใจง่าย หลักการที่ใช้ในการแยกส่วนภาพมี 2 หลักการ หลักการแรกคือ แยกส่วนภาพจากความไม่ต่อเนื่อง (Discontinuity) ขององค์ประกอบภาพ โดยดูจากความเปลี่ยนแปลงอย่างชัดเจนของความเข้มขององค์ประกอบภาพ เช่น บริเวณที่เป็นเส้นขอบของภาพ หลักการที่สองคือ การแยกส่วนภาพตามความคล้ายกัน (Similarity) ขององค์ประกอบภาพ โดยแยกส่วนภาพตามคุณสมบัติของจุดภาพ (Pixel) ภายในพื้นที่เดียวกันที่มีความเหมือนกัน เทคนิคที่ใช้ในการแยกส่วนภาพมีดังต่อไปนี้

- การตรวจจับเส้นขอบของภาพ (Edge Detection) เส้นขอบ (Edge) เป็นชุดของจุดภาพที่เชื่อมต่อกันวางตัวบนขอบระหว่าง 2 พื้นที่ (Region) ที่มีค่าระดับเทาต่างกัน แบบจำลองของเส้นขอบหาได้จากการเปลี่ยนค่าระดับเทา (Gray Levels) ของจุดภาพมีลักษณะของแบบจำลองเป็นรูปลาดเอียง (Ramp Like)
- การหาขีดเริ่มเปลี่ยน (Thresholding) ของภาพเป็นกระบวนการในการสร้างพื้นที่ที่มีความเป็นเอกกรูป (Uniformity) ในรูปภาพออกเป็นส่วนๆ ตามเกณฑ์ของขีดเริ่มเปลี่ยน (Threshold Criterion: T) ซึ่งแสดงดังสมการ 1 เมื่อ T เป็นฟังก์ชันของเกณฑ์ของขีดเริ่มเปลี่ยน $f(x,y)$ เป็นค่าระดับเทาของจุดภาพที่ (x,y) และ $A(x,y)$ แทนคุณสมบัติของจุดภาพข้างเคียง

$$T = T\{x, y, A(x, y), f(x, y)\} \quad (1)$$

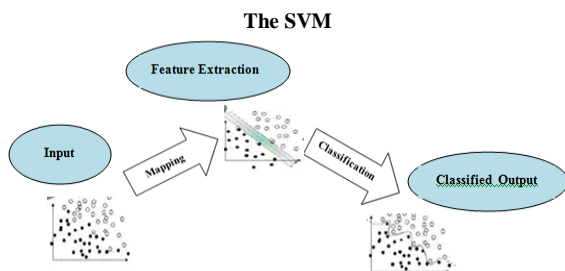
ภาพที่ทำการหาขีดเริ่มเปลี่ยนแล้ว $g(x,y)$ มีค่าดังสมการที่ 2

$$g(x, y) = \begin{cases} 1 & \text{ถ้า } f(x, y) \leq T \\ 0 & \end{cases} \quad (2)$$

อัลกอริทึมที่ใช้ในการตรวจหาเส้นขอบของภาพที่ใช้ในปัจจุบันมีดังนี้ วิธีของโซเบล (Sobel method) หรือ SB [6], [17] วิธีของพรีวิตต์ (Prewitt method) หรือ PW [7], [17] วิธีของโรเบิร์ตส์ (Roberts method) หรือ RB [8], [17] วิธีลาปลาเซียนของเกาส์ (Laplacian of Gaussian method) หรือ LOG วิธีตัดขวางศูนย์ (Zerocross method) หรือ ZC [9], [17] วิธีของแคนนี่ (Canny method) หรือ CN [10], [17] เป็นต้น

ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine : SVM)

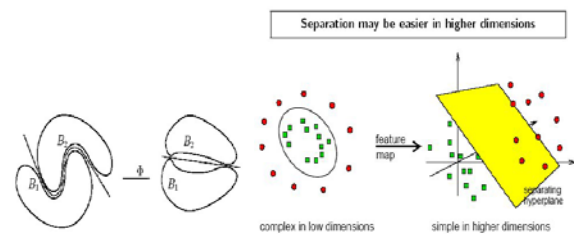
เป้าหมายของวิธีซัพพอร์ตเวกเตอร์แมชชีน คือ กระบวนการสอนเครื่องแบบมีผู้สอน (Supervise Learning) เพื่อให้สามารถสร้างตัวจัดประเภทข้อมูล(Classifier) ที่มีความเป็นทั่วไป (Generalize) สูง นั่นคือ สามารถทำงานได้ดีกับตัวอย่างที่ไม่รู้จัก (Unknown Dataset) ด้วยกระบวนการปรับรูปแบบข้อมูลจากข้อมูลที่มีมิติต่ำ (Low Dimension Dataset) บนพื้นที่ข้อมูลนำเข้า (Input Space) ให้อยู่ในรูปแบบของข้อมูลที่มีมิติสูง (High Dimension Dataset) บนพื้นที่ข้อมูลคุณลักษณะ (Feature Space) โดยใช้ฟังก์ชันในการปรับรูปแบบข้อมูลที่เรียกว่า ฟังก์ชันเคอร์เนล (Kernel Function) ซึ่งความสามารถดังกล่าวช่วยให้การสร้างตัวจัดประเภทข้อมูลด้วยสมการกำลังสอง (Quadratic Equation) บนพื้นที่ข้อมูลคุณลักษณะเป็นไปได้ง่ายขึ้น และมีความชัดเจนในการจัดประเภทมากยิ่งขึ้นด้วย นอกจากนี้ ตัวจัดประเภทข้อมูลที่ตีควรมีโครงสร้างแบบเส้นตรง (Linear Classifier) และสามารถสร้างพื้นที่ระยะห่างระหว่างตัวจัดประเภทข้อมูลเองกับค่าที่ใกล้ที่สุดของแต่ละกลุ่มข้อมูลได้มากที่สุด เพื่อประสิทธิภาพในการแยกแยะประเภทของชุดข้อมูลแต่ละประเภทออกจากกันอย่างชัดเจน ซึ่งเส้นที่เหมาะสมดังกล่าวถูกเรียกว่า ระนาบแบ่งเขตข้อมูลที่เหมาะสม (The Optimal Separating Hyper plane) โดยหลักการในการทำงานเพื่อจำแนกประเภทข้อมูลของวิธีซัพพอร์ตเวกเตอร์แมชชีนสามารถแสดงตามภาพที่ 1



ภาพที่ 1: อัลกอริทึมของ Support Vector Machines

ฟังก์ชันเคอร์เนล (Kernel Function) ในกรณีการจัดแบ่งกลุ่มข้อมูล โดยใช้ระนาบแบบไม่เป็นเส้นตรงซัพพอร์ตเวกเตอร์แมชชีนจะอาศัยหลักการของการแปลงข้อมูลจากพื้นที่ข้อมูลนำเข้า (Input Space) ให้เป็นพื้นที่คุณลักษณะ (Feature Space) ที่มีมิติสูงขึ้น จากภาพที่ 1 และภาพที่ 2 แสดงให้เห็นถึงแนวคิดของซัพพอร์ตเวกเตอร์แมชชีน ซึ่งทำการแปลงข้อมูลที่ใช้

ในการเรียนรู้แบบไม่เชิงเส้น ไปเป็นขนาดพื้นที่คุณลักษณะที่ใหญ่ขึ้นผ่านฟังก์ชันเคอร์เนล (Kernel Function : Φ) และสร้างระนาบซึ่งแบ่งข้อมูลสองกลุ่มได้ดีที่สุด ทำให้เกิดเป็นขอบเขตการตัดสินใจ (Decision Surface) แบบไม่เชิงเส้นในพื้นที่ข้อมูลนำเข้า ในขณะที่ซัพพอร์ตเวกเตอร์แมชชีนแบบเชิงเส้นจะสร้างระนาบในพื้นที่คุณลักษณะที่ใหญ่ขึ้นภายใต้ทฤษฎีของ Mercer (Courant and Hilbert, 1953)[9] ซึ่งต้องการการคำนวณที่สิ้นเปลืองในส่วนของตัวอย่างเพื่อให้ได้ขนาดพื้นที่คุณลักษณะที่ใหญ่ขึ้น ปัญหาดังกล่าวสามารถแก้ไขได้โดยใช้ฟังก์ชันเคอร์เนล เพื่อให้ได้ผลลัพธ์ที่น่าพึงพอใจเช่นเดียวกัน ซึ่งการใช้ฟังก์ชันเคอร์เนลจะทำให้สามารถคำนวณระนาบได้โดยไม่ต้องอาศัยการแปลงไปเป็นพื้นที่คุณลักษณะ



ภาพที่ 2: แนวความคิดการจำแนกประเภทข้อมูล

ของวิธีซัพพอร์ตเวกเตอร์แมชชีน

ในปัจจุบันมีงานวิจัยที่เกี่ยวข้องในการตรวจหาจุดบกพร่องของพื้นผิวแบ่งออกได้หลายกลุ่ม อาทิ [11] ทำการศึกษาตรวจจับจุดบกพร่องบนผ้าทอโดยใช้ DWT และดึงค่าพลังงานสูงสุดโดยใช้GLCM (Gray Level Co- Occurrence Matrix) ซึ่งพบว่าวิธีการดังกล่าวได้ผลประสิทธิภาพ 94% ส่วน [12] ได้ทำการศึกษาโดยในการดึงลักษณะใช้การปรับค่า Gray image ไปเป็น Binary image และทำการแยกด้วย BPNN (Black Propagation Neural Network) ผลการศึกษาพบว่า ได้ผลประสิทธิภาพ 98.56% สำหรับ [13] เสนอการใช้ Statistical feature ดึงคุณลักษณะ และใช้ NN (Neural Network) ในการแยกประเภท พบว่า วิธีการที่น่าเสนอได้ผลประสิทธิภาพ 98% [14] ทำการศึกษาการตรวจจับจุดบกพร่องบนพื้นผิวผ้าในอุตสาหกรรมผ้าทอ โดยใช้ Wavelet ร่วมกับ GA (Genetic Algorithm) ซึ่งวิธีดังกล่าวเป็นวิธีที่มีประสิทธิภาพในการตรวจจับจุดบกพร่องบนพื้นผิวผ้าทอ และลดเวลาในการ

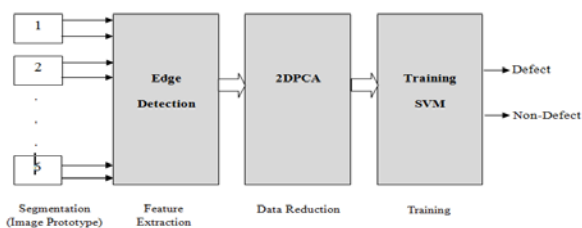
ประมวลผลผลง นอกจากนี้ [15] ได้ทำการศึกษาการตรวจจับจุดบกพร่องบนพื้นผิวผ้า โดยใช้ Undecimated Wavelet Transform ในการดึงคุณลักษณะ [16] ได้ศึกษาและนำเสนอการดึงคุณลักษณะด้วย Adaptive Quantized Gray-level Co- Occurrence Matrix ช่วยลดขนาดและในการหาคุณลักษณะใช้ SVDD (Support Vector Description Data) และ [17] นำเสนอวิธีในการดึงคุณลักษณะโดยใช้การหาเส้นขอบภาพ (Edge Detection) และการวิเคราะห์หึ่งค์ประกอบหลักสองมิติ (2DPCA) ร่วมกับ GA (Genetic Algorithm) เพื่อลดข้อมูล และใช้ ANN (Artificial Neural Network) ในการแยกจุดบกพร่อง ซึ่งพบว่ารูปแบบที่นำเสนอข้างต้นเป็นวิธีที่มีประสิทธิภาพในการตรวจจับจุดบกพร่อง แต่ใช้เวลามากในการประมวลผล จะเห็นได้ว่ารูปแบบที่นำเสนอโดยนักวิจัยหลายท่านให้ผลประสิทธิภาพในการตรวจจับจุดบกพร่องบนพื้นผิวที่แตกต่างกันไป และมีข้อจำกัดบางประการที่น่าสนใจ จึงมีการวิจัยและพัฒนาอัลกอริทึมสำหรับกระบวนการตรวจสอบจุดบกพร่องมาอย่างต่อเนื่อง มีการคิดค้นเทคนิคและวิธีการใหม่ๆ เพื่อให้เกิดประสิทธิภาพและความแม่นยำอยู่เสมอ ในงานวิจัยนี้สนใจที่จะนำเสนอการประยุกต์ใช้การหาเส้นขอบภาพ (Edge detection) ในการดึงคุณลักษณะ ในการลดข้อมูลใช้การวิเคราะห์หึ่งค์ประกอบหลักสองมิติ (2DPCA) และซัพพอร์ตเวกเตอร์แมชชีน (SVM) สำหรับการตรวจจับจุดบกพร่องบนสิ่งทอเพื่อให้ได้ความถูกต้องแม่นยำที่สูงขึ้น เพื่อให้สามารถตรวจจับจุดบกพร่องบนสิ่งทอได้อย่างมีประสิทธิภาพและครอบคลุมปัจจัยที่เกี่ยวข้องให้มากที่สุด

3. วิธีการดำเนินการวิจัย

ในการดำเนินการวิจัย แบ่งออกเป็น 2 ขั้นตอน ดังนี้

1) การฝึกสอนเพื่อให้ได้ผลลัพธ์ที่จะนำไปใช้เวลาจริง

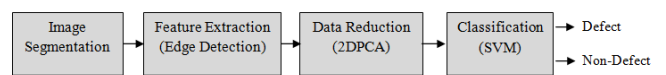
มีการดำเนินการ ดังภาพที่ 3



ภาพที่ 3: การฝึกสอน

ใช้ภาพขนาด 100x100x3 พิกเซล เป็นขนาดของภาพต้นแบบ ต้นแบบเป็นภาพมีจุดบกพร่อง 5 แบบ และไม่มีจุดบกพร่อง 1 แบบ แต่ละแบบใช้จำนวน 20 ภาพ รวมเป็นจำนวนภาพทั้งสิ้น 120 ภาพ เปลี่ยนเป็นค่าระดับสีเทา นำไปทำการดึงลักษณะเด่นด้วยการใช้การหาเส้นขอบภาพ (Edge detection) โดยวิธีลาปลาเซียนของเกาส์ (Laplacian of Gaussian method) หรือ LOG แล้วทำการลดข้อมูลให้มีขนาดเล็กลง ด้วยการวิเคราะห์หึ่งค์ประกอบหลักสองมิติ (2DPCA) จากนั้นนำไปทำการฝึกสอนโดยอัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน (SVM) จนได้ค่าที่ให้ผลลัพธ์ตามที่ต้องการ และจะนำผลลัพธ์ที่ได้จากการทำการฝึกสอนแล้วไปใช้งานขณะทำงานจริงต่อไป

2) การทดสอบใช้งานเวลาจริง ดังภาพที่ 4



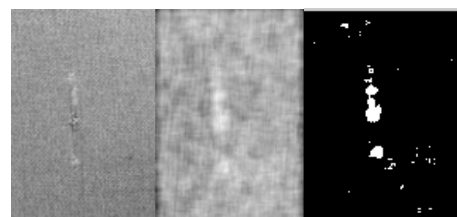
ภาพที่ 4: การใช้งานขณะประมวลผลเวลาจริง

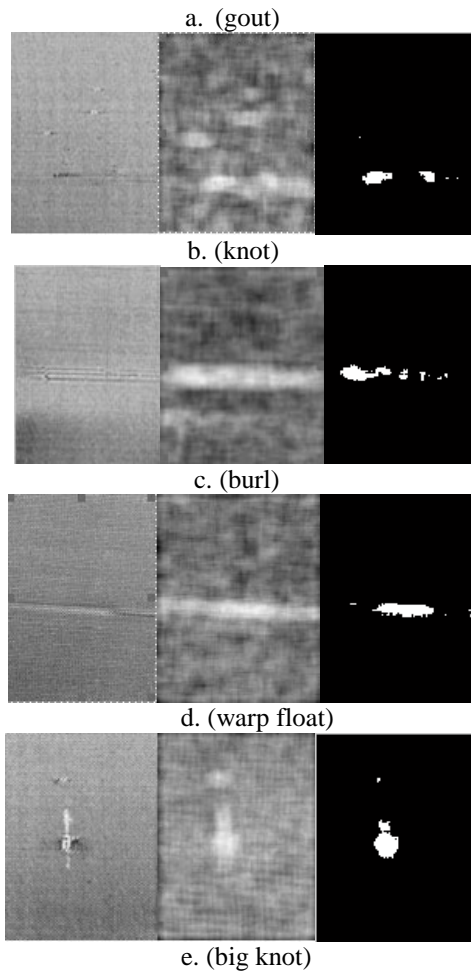
จับภาพที่เวลาจริงด้วยขนาด 100x100x3 พิกเซล เปลี่ยนเป็นค่าระดับเทาผ่านสู่การหาเส้นขอบภาพ (Edge detection) โดยวิธีลาปลาเซียนของเกาส์ (Laplacian of Gaussian method) หรือ LOG ที่ได้จากขั้นตอนที่ 1 ลดขนาดด้วยการวิเคราะห์หึ่งค์ประกอบหลักสองมิติ (2DPCA) เช่นเดียวกันด้วยผลลัพธ์จากขั้นตอนที่ 1 ทำการประมวลผลเก็บสถิติ เพื่อทำการวิเคราะห์ต่อไป

4. ผลการทดลอง

ในงานวิจัยนี้ได้ทำการทดสอบกับภาพทั้งหมด 120 ภาพ ซึ่งมีขนาดภาพเริ่มแรกเป็น 100x100x3 พิกเซล โดยการกำหนดขีดเริ่มเปลี่ยน (Threshold) มีค่าเท่ากับ 0.8 ภาพที่ได้เป็นภาพต้นแบบของจุดบกพร่องทั้ง 5 แบบ แสดงดังภาพที่ 5

Image Fabric Image 2Gray Edge Detection



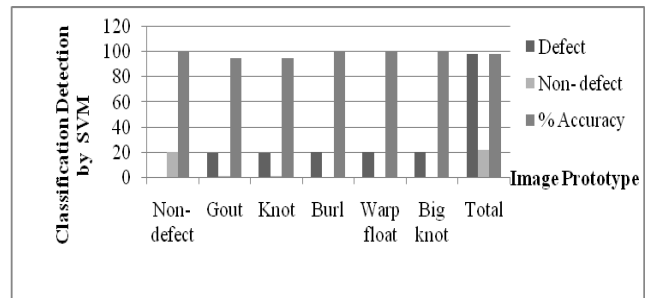


ภาพที่ 5: ภาพตัวอย่างจากการตรวจจับจุดบกพร่อง

ตารางที่ 1 ผลการตรวจจับจุดบกพร่องอัตโนมัติบนสิ่งทอ

Image Prototype	Classification Detection by SVM		
	Defect	Non- defect	% Accuracy
Non- defect	0	20	100
Gout	19	1	95
Knot	19	1	95
Burl	20	0	100
Warp float	20	0	100
Big knot	20	0	100
Total	98	22	98.33

จากตารางที่ 1 แสดงประสิทธิภาพการตรวจจับจุดบกพร่องอัตโนมัติบนสิ่งทอ โดยการผสมผสานระหว่างการหาเส้นขอบภาพร่วมกับซัพพอร์ตเวกเตอร์แมชชีน จะเห็นว่า วิธีการดังกล่าวสามารถตรวจจับจุดบกพร่องบนสิ่งทอที่ความแม่นยำสูงสุดที่ 98.33 เปอร์เซ็นต์ แสดงดังภาพที่ 6



ภาพที่ 6: กราฟความสัมพันธ์ระหว่างค่าความแม่นยำ (Accuracy) กับจุดบกพร่องแบบต่างๆ

กราฟจากภาพที่ 6 แสดงความสัมพันธ์ระหว่างค่าความแม่นยำ (Accuracy) กับจุดบกพร่องแบบต่างๆ จากการตรวจจับจุดบกพร่องด้วยการหาเส้นขอบภาพร่วมกับซัพพอร์ตเวกเตอร์แมชชีน สามารถอธิบายได้ดังนี้ เมื่อพิจารณาแต่ละแบบของจุดบกพร่อง/ไม่บกพร่อง พบว่า ค่าความแม่นยำของการทดสอบการตรวจจับจุดบกพร่องบนสิ่งทอ เมื่อใช้วิธีการดังกล่าวโดย จุดบกพร่องแบบก้อนเส้นไหม (Burl), แบบปมเส้นไหมตามยาว (Warp float), แบบกระจุกใหญ่ (Big Knot) และแบบไม่มีจุดบกพร่อง (Non-defect) มีค่าความแม่นยำมากที่สุด คิดเป็น 100% รองลงมา คือ จุดบกพร่องแบบปมผ้า (Gout) และแบบกระจุก (Knot) คิดเป็น 95% และจากผลการดำเนินงานในการศึกษาและทดลองทั้งหมดข้างต้น พบว่า วิธีการที่นำเสนอดังกล่าว สามารถตรวจจับจุดบกพร่องบนสิ่งทอได้ด้วยความแม่นยำสูงสุดคิดเป็น 98.33 %

5. สรุป

จากการศึกษาผลการตรวจจับจุดบกพร่องด้วยการผสมผสานระหว่างการหาเส้นขอบภาพร่วมกับซัพพอร์ตเวกเตอร์แมชชีน พบว่า รูปแบบที่นำเสนอเป็นวิธีที่มีประสิทธิภาพที่ดีในการตรวจจับจุดบกพร่องบนสิ่งทอ ใช้เวลาน้อยในการฝึกสอนอย่างไรก็ตามวิธีดังกล่าว ยังมีข้อจำกัดในเรื่องสภาวะแสงที่แตกต่างกัน กล่าวคือ สามารถใช้ได้ ในสภาวะแสงที่เหมาะสมในตอนกลางวันเท่านั้น และต้องมีการกำหนดระยะจุดโฟกัสในการตรวจจับจุดบกพร่องที่เหมาะสม นอกจากนี้แล้วรูปแบบที่นำเสนอที่สามารถใช้งานได้อย่างหลากหลาย จำเป็นต้องมีข้อมูลที่ใช้ในการฝึกสอนที่หลากหลาย และมีจำนวนเพียงพอที่จะ

ใช้แทนกลุ่มข้อมูลได้ทุกกลุ่ม งานวิจัยนี้เป็นงานวิจัยเริ่มต้นเท่านั้น ดังนั้น ในงานวิจัยต่อไปควร มีการพัฒนาและนำไปประยุกต์ใช้ในงานอื่นๆต่อไป

เอกสารอ้างอิง

- [1] R.Chin, "Automated Visual Inspection Techniques and Applications : A Bibliography," *Pattern Recognition*15(4), pp.343-357,1982.
- [2] R. Gonzalez and R. Woods., "Digital Image Processing," 2nd edn. *Addison-Wesley Pub-lishing Company*, Reading, 2002.
- [3] R. Mehrotra , K. Namuduri and N. Ranganathan, "Gabor filter-based edge detection," *Pattern Recognition* 25, pp. 1479-1494, 1992.
- [4] D.,Cesacent and J. Smokelin, "Neural net design of Gabor wavelet filters for distortion invariant object detection in cluster," *Optical Engineering* 33, pp. 2264-2271, 1994.
- [5] G. Peng, "Image processing : interpolation," 2004.
- [6] A. Srikaew., "Genetic Algorithms-Part I," *Suranaree Journal of Science and Technology*, vol. 9, no. 1, pp 69-83, Jan-Mach 2002.
- [7] J. Yang, D. Zhang and A. Frangi, "Two-dimensional PCA : A new approach appearance-based face representation and recognition,". *IEEE Trans.On Pattern Analysis and Machin Intelligence*, vol. 6, no. 2.2, pp. 131-137, 2004.
- [8] J. Yang, D. Zhang, A. Frangi, and J. Yang., "Tow dimensional PCA : A new approach to appearance-based face representation and recognition," *IEEE Transactions on pattern analysis and machine intelligence*, no. 26, pp. 131-137, 2004.
- [9] Y. Ben Salem., "Automatic recognition of fabrics based on texture and using SVM," *Springer : Original paper*. 2010.
- [10] A. Srikaew, K. Attakitmoncol and W. Kidsang., "Detection of Defect in Textile Fabrics using Optimal Gabor Wavelet Network and Two-Dimonsional PCA," *Suranaree Journal of Science and Technology*, Vol. 18,no.1, pp 436-445, 2011.
- [11] N. Dhivya, R. Anusuya Lakshmi and D. Ramalakshmi, "A Hybrid Approach for Detection And Classification of Fabric Faults In Textile Industries Using Wavelet Transform And GLCM Technique," *International Journal of Advanced Electrical and Electronics Engineering, (IAEEEE)*, ISSN (Print): 2278-8948, Vol.2, Issue-4, 2013.
- [12] A. Chakraborty, P. Kanti Kumar and S.Moni Chatterjee, "Detection of Lousiness in Silk Fabric Using Digital Image Processing," *The 7th International Conference - TEXSCI 2010*, Liberec, Czech Republic, Sep 6-8,2010.
- [13] Md. Tarek Habib and M. Rokonzuzaman, "Distinguishing Feature Selection for Fabric Defect Classification Using Neural Network," *Journal of Multimedia*, Vol. 6, No. 5, pp.416-424, 2011.
- [14] N. Heidari, R. Azmi and B. Pishgoo, "Fabric Textile Defect Detection, By Selecting A Suitable Subset Of Wavelet Coefficients, Through Genetic Algorithm," *International Journal of Image Processing (IJIP)*, Vol.5, No1, pp.25-35, 2011.
- [15] B. Mingde, S. Zhigang, "Fabric Defect Detection using Undecimated Wavelet Tranfrom," *Information Technology Journal*, Vol.10, No.9, pp.1701-1708, 2011.
- [16] B. Mingde, S. Zhigang and L. Yesong, "Textural Fabric Defect Detection using Adaptive Quantized Gray-level Co-cocurrence Matrix and Support Vector Description Data," *Information Technology Journal*, Vol.11, No.6, pp.673-685, 2012.
- [17] T. Prachpreecha and K. Kubola. "Defect Detection in Textile Fabrics with Edge Detection and Artificial Neural Network," *Proceedings of the 8th National Conference on Computing and Information Technology (NCCIT12)*, Pattaya, Chonburi,Thailand, May 9-10,2012.
- [18] K. Ajay, and Grantham K. H. Pang, "Defect Detection in Textured Materials Using Gabor Filter," *IEEE Trans.On Industry Applications*, vol. 36, no. 2.2, March/April 2002.
- [19] P.K Mensah and P.K Buah-Bassuah, "Surface Defect Detection for Some Ghanaian Textile Fabrics using Moire Interferometry," *Research Journal of Applied Sciences, Engineering and Technology* , Vol.6, No.3, pp.349-353, 2013.
- [20] วิไลลักษณ์ ทิดสร้าง และ อาทิตย์ ศรีแก้ว, "การตรวจจับจุดบกพร่องด้วยตัวกรองกอบอร์" วารสารวิชาการ: การสัมมนาทางวิชาการบัณฑิตศึกษาวิศวกรรมไฟฟ้า ประจำปีครั้งที่ 4 มหาวิทยาลัยเทคโนโลยีสุรนารี. มีนาคม 2550 หน้า 73-80.

การศึกษาเปรียบเทียบวิธีการหาลักษณะเด่นบนอุปกรณ์วัดโครงสร้างมือแบบมีหมุดกำกับ

A Comparative Study in Feature Extraction Methods of Peg-fixed Hand Geometry Device

จักรี ศรีนนท์ฉัตร (Jakkree Srinonchat) และอภิวัฒน์ สวัสดิ์ศิริรัตน์ (Apiwat Sawatdirat)

ภาควิชาวิศวกรรมอิเล็กทรอนิกส์และโทรคมนาคม คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
jakkree.s@en.rmutt.ac.th, apiwat.sa@rmuti.ac.th

บทคัดย่อ

กระบวนการหาลักษณะเด่นของอุปกรณ์วัดโครงสร้างมือเป็นปัจจัยหลักที่ส่งผลกระทบต่อประสิทธิภาพของระบบวัดโครงสร้างมือ งานวิจัยนี้จึงมุ่งเน้นที่การเปรียบเทียบเทคนิคการหาลักษณะเด่นบนอุปกรณ์วัดโครงสร้างมือแบบมีหมุดกำกับ ระหว่างวิธีการนับพิกเซลจากรูปเงามือกับวิธีการคำนวณจากเส้นรอบรูปมือ เพื่อเป็นแนวทางในการนำมาประยุกต์ใช้บนระบบสมองกลฝังตัว ผลการทดลองสรุปได้ว่าวิธีการนับพิกเซลจากรูปเงามือมีข้อดีด้านความเร็วในการประมวลผล ส่วนวิธีการคำนวณจากเส้นรอบรูปมือมีข้อดีในด้านความยืดหยุ่นสูงสามารถรับข้อมูลที่มีขนาดแตกต่างกันได้ดีกว่าและสามารถพัฒนาสู่ระบบวัดโครงสร้างมือแบบไร้หมุดกำกับได้

คำสำคัญ: อุปกรณ์วัดโครงสร้างมือ ระบบสมองกลฝังตัว การหาลักษณะเด่น เส้นรอบรูปมือ คอนเวกซ์ฮัลล์

Abstract

Feature extraction process of hand geometry device is a majorly effected to the efficiency of hand geometry system. This research presents a comparative study in feature extraction methods of peg-fixed hand geometry device between the counting pixel-based on hand silhouette technique and the hand contour-based technique. This technique can be applied to the embedded system. The experiment results show that the counting pixel-based on hand silhouette technique offers the fastest processing as the advantage. While the hand contour-based technique offers the advantage of using with many different inputs which can be developed to peg-free hand geometry system.

Keywords: Hand Geometry Device, Embedded System, Feature Extraction, Hand Contour, Convex Hull.

1. บทนำ

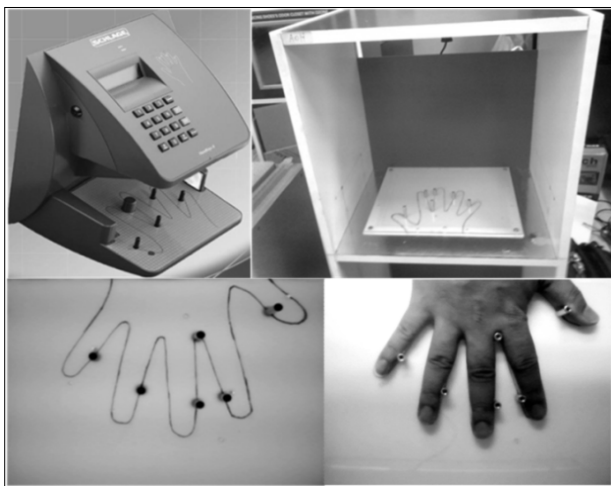
ระบบวัดโครงสร้างมือ (Hand Geometry System) [1,2] คือระบบไบโอเมตริกที่นำข้อมูลขนาดและรูปร่างของฝ่ามือมาจำแนกความแตกต่างของแต่ละบุคคลออกจากกัน ระบบนี้มีข้อดีด้านความรวดเร็วในการประมวลผล ใช้พื้นที่ในการจัดเก็บน้อยเนื่องจากข้อมูลมีขนาดเล็ก จึงเหมาะกับอุปกรณ์ที่มีพื้นที่เก็บข้อมูลจำกัดหรือประยุกต์ใช้ในรูปแบบเครือข่าย สะดวกในการใช้งานและเป็นที่ยอมรับของสาธารณะ แต่เนื่องจากระบบที่มีวางจำหน่ายมีราคาแพงเนื่องจากมีเพียงไม่กี่บริษัทที่พัฒนาระบบในเชิงพาณิชย์ ผู้วิจัยจึงได้มีแนวคิดที่จะทำการพัฒนาต้นแบบระบบระบบวัดโครงสร้างมือเพื่อการยืนยันเอกลักษณ์ (Verification) ด้วยเทคโนโลยีสมองกลฝังตัวในแบบระบบเปิดเพื่อเป็นอีกทางเลือกนอกเหนือจากระบบในเชิงพาณิชย์ โดยงานวิจัยนี้จะมุ่งเน้นไปที่การเปรียบเทียบเทคนิคการหาลักษณะเด่นบนอุปกรณ์วัดโครงสร้างมือแบบมีหมุดกำกับนิ้วมือ (Peg-fixed) ระหว่างวิธีการนับพิกเซลจากรูปเงามือ (Counting pixel-based) ดังงานวิจัย [3,4] กับวิธีการคำนวณจากเส้นรอบรูปมือ (Hand contour-based) ดังงานวิจัย [5-7] โดยจะเปรียบเทียบในด้านความเร็วในการประมวลผล ความเที่ยงตรง พร้อมทั้งหาปัจจัยที่มีผลกระทบกับความถูกต้องของวิธีการ

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 อุปกรณ์วัดโครงสร้างมือ

อุปกรณ์วัดโครงสร้างมือเป็นส่วนหนึ่งของระบบวัดโครงสร้างมือ ทำหน้าที่รับภาพมือของผู้ใช้และหาลักษณะเด่น

ออกมาเพื่อส่งต่อไปยังกระบวนการขั้นต้นเอกลักษณ์มีขั้นตอนการทำงานหลักแบ่งเป็นสามส่วนคือ การเตรียมภาพ (Image Acquisition), การประมวลผลภาพเบื้องต้น (Image Pre-processing) และการหาลักษณะเด่น (Feature Extraction) ในส่วนของการเตรียมภาพ งานวิจัยนี้ได้ใช้กล้องเว็บแคม (Web Camera) บันทึกภาพมือขวาจากตู้ควบคุมแสงสว่าง ซึ่งมีหมุด 6 หมุดช่วยบังคับมือ ให้อยู่ในตำแหน่งที่ต้องการ มีความชัดเจนของนิ้วมือทั้ง 4 นิ้วคือ นิ้วก้อย, นิ้วนาง, นิ้วกลางและนิ้วชี้ คล้ายกับผลิตภัณฑ์เชิงพาณิชย์ดังภาพที่ 1 และงานวิจัย [3-10] ในส่วนการประมวลผลภาพเบื้องต้นจะนำภาพมือของผู้ใช้มาปรับปรุงให้เหลือเฉพาะส่วนที่ต้องการ แล้วแปลงเป็นภาพขาวดำ ส่วนสุดท้ายคือการหาลักษณะเด่น ขั้นตอนนี้จะนำภาพที่ได้มาหาลักษณะเด่นของมือในแต่ละบุคคลออกมาในรูปแบบของชุดข้อมูลเพื่อประหยัดพื้นที่ในการบันทึกและความสะดวกในการนำไปเปรียบเทียบ เช่น ความกว้างและความยาวของแต่ละนิ้ว, อัตราส่วนระหว่างนิ้วกับฝ่ามือ และความหนาของมือ เป็นต้น ซึ่งจะแตกต่างกันไปในแต่ละงานวิจัยนั้นๆว่าต้องการความถูกต้องแม่นยำหรือความผิดพลาดตั้งไว้ในระดับใด

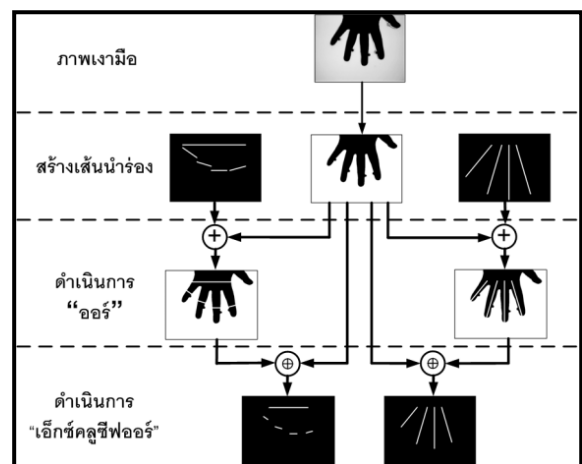


ภาพที่ 1: (ซ้ายบน) ผลิตภัณฑ์เชิงพาณิชย์ (ขวาบน) ตู้ควบคุมแสง (ซ้ายล่าง) แผ่นรองมือพร้อมหมุด (ขวาล่าง) ตัวอย่างภาพที่ได้

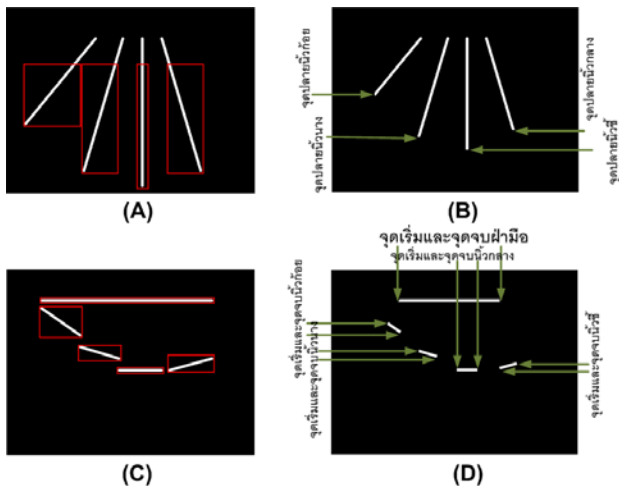
2.2 วิธีการนับพิกเซลจากรูปเงามือ

กระบวนการหาลักษณะเด่นของมือด้วยวิธีการนับพิกเซลจากรูปเงามือเพื่อหาความยาวและความกว้างของนิ้วทั้งสี่นิ้วคือนิ้วก้อย, นิ้วนาง, นิ้วกลาง และนิ้วชี้ พร้อมทั้งหาความกว้างของฝ่ามือ รวมทั้งสิ้นเก้าค่า มีการทำงานสี่ขั้นตอนคือ ขั้นตอน

หนึ่งสร้างภาพเส้นนำร่องแบบคงที่ (Fixed Guide Lines) สองภาพคือภาพเส้นนำร่องความยาว และภาพเส้นนำร่องความกว้าง จากตำแหน่งของหมุดกำกับฝ่ามือเก็บไว้ในระบบ ซึ่งเส้นเหล่านี้จะเป็นตัวจำกัดความยาวและความกว้างสูงสุดที่ระบบจะวัดได้ ขั้นตอนที่สองนำภาพเส้นนำร่องทั้งสองมาดำเนินการทางลอจิกออร์กับภาพมือขาวดำ ได้ผลลัพธ์เป็นภาพเงามือพร้อมเส้นนำร่องความยาวนิ้วมือและภาพเงามือพร้อมเส้นนำร่องความกว้างนิ้วมือ ขั้นตอนที่สามนำภาพผลลัพธ์จากขั้นตอนที่สองมาดำเนินการทางลอจิกเอ็กซ์คลูซีฟออร์กับภาพมือขาวดำ จะได้ภาพเส้นผลลัพธ์ความยาวนิ้วมือและภาพเส้นผลลัพธ์ความกว้างนิ้วมือ ดังภาพที่ 2 ในส่วนของขั้นตอนสุดท้ายคือการหาความยาวนิ้วและความกว้างนิ้ว ดังภาพที่ 3 โดยการหาความยาวนิ้วมือต้องทำการกำหนดกรอบการค้นหา เพื่อความสะดวกในการระบุนิ้วมือและลดเวลาในการประมวลผลดังภาพที่ 3(A) จากนั้นค้นหาจุดปลายนิ้วซึ่งหาได้จากพิกเซลสีขาวตำแหน่งแรกจากกรอบที่กำหนด โดยเริ่มค้นหาจากกรอบที่มีค่าพิกเซลมากที่สุดลงมาหาค่าพิกเซลที่น้อยที่สุดดังภาพที่ 3(B) และคำนวณหาความยาวของนิ้วด้วยสมการที่ 1 ส่วนการหาความกว้างนิ้วมือและความกว้างฝ่ามือก็เช่นเดียวกันต้องทำการกำหนดกรอบการค้นหา ดังภาพที่ 3(C) แต่ต้องค้นหาจุดเริ่มนิ้วซึ่งหาได้จากพิกเซลสีขาวตำแหน่งแรกจากกรอบที่กำหนด โดยเริ่มค้นหาจากกรอบที่มีค่าพิกเซลน้อยที่สุดขึ้นไปหาค่าพิกเซลที่มากที่สุด และค้นหาจุดจบนิ้วซึ่งหาได้จากพิกเซลสีขาวตำแหน่งแรกจากกรอบที่กำหนด โดยเริ่มค้นหาจากกรอบที่มีค่าพิกเซลมากที่สุดลงมาหาค่าพิกเซลที่น้อยที่สุดดังภาพที่ 3(D) แล้วนำจุดทั้งสองที่ได้มาคำนวณหาความกว้างของนิ้วด้วยสมการที่ 1 เช่นกัน



ภาพที่ 2: วิธีการนับพิกเซลจากรูปเงามือ



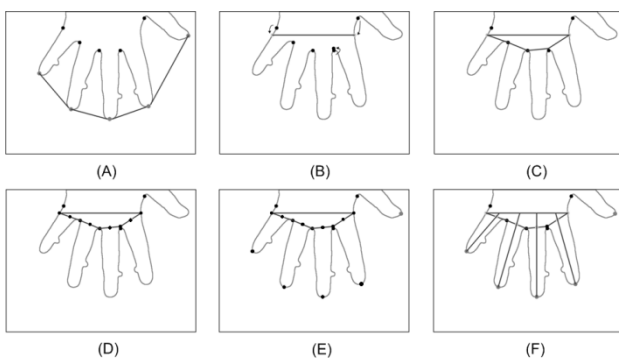
ภาพที่ 3: การหาความยาวนิ้วและความกว้างนิ้ว

$$\text{length} = \sqrt{|X_{\text{start}} - X_{\text{end}}|^2 + |Y_{\text{start}} - Y_{\text{end}}|^2} \quad (1)$$

เมื่อ X คือ พิกัดในแนวแกนอนในระบบพิกัดฉาก
 Y คือ พิกัดในแนวแกนตั้งในระบบพิกัดฉาก
 start คือ จุดเริ่ม
 end คือ จุดจบ

2.3 วิธีการคำนวณจากเส้นรอบรูปมือ

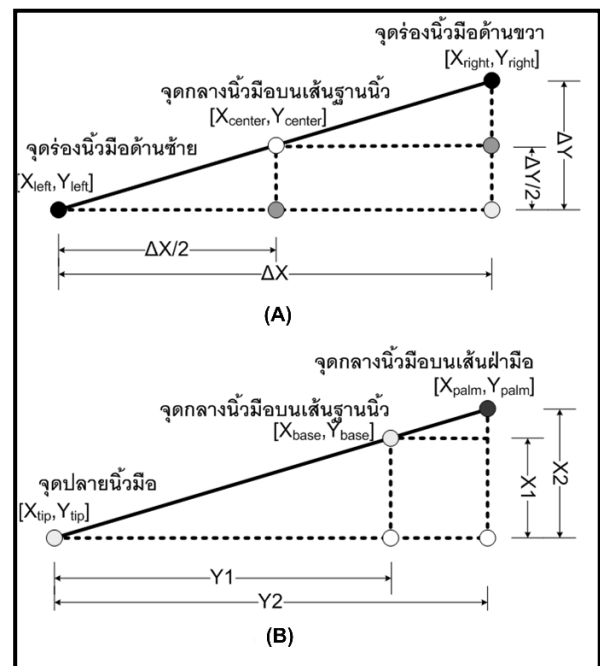
กระบวนการหลักขณะเด่นของมือด้วยวิธีการคำนวณจากเส้นรอบรูปมือเพื่อหาความยาวนิ้วและความกว้างของฐานนิ้ว ทั้งสี่นิ้วคือ นิ้วก้อย, นิ้วนาง, นิ้วกลาง และนิ้วชี้ พร้อมทั้งหาความกว้างของฝ่ามือ รวมทั้งสิ้นเก้าค่า ในงานวิจัยนี้แบ่งเป็นหกขั้นตอน ดังแสดงในภาพที่ 4 ดังนี้



ภาพที่ 4: วิธีการคำนวณจากเส้นรอบรูปมือ

ภาพที่ 4(A) แสดงการหาปลายนิ้วและร่องนิ้ว ซึ่งในงานวิจัยนี้ใช้การหาปลายนิ้วและร่องนิ้วด้วยวิธีการหากรอบด้านนอกและความเว้า (Convex hull and Convexity Defects)

ของเส้นรอบรูปมือจากงานวิจัย [11-12] ภาพที่ 4(B) แสดงการปรับตำแหน่งข้อมือทั้งซ้ายและขวาเพื่อใช้เป็นจุดอ้างอิงในการหาความกว้างของฐานนิ้วและความกว้างของฝ่ามือ และปรับตำแหน่งร่องนิ้วระหว่างนิ้วกลางกับนิ้วชี้ซึ่งมีมุมก้ำก้อยเพื่อให้ได้ผลลัพธ์ที่ถูกต้องขึ้น ภาพที่ 4(C) แสดงการหาความกว้างของฐานนิ้วมือ และความกว้างของฝ่ามือ โดยในส่วนของความกว้างของฐานนิ้วมือแต่ละนิ้วจะคำนวณจากตำแหน่งร่องนิ้วที่มีลำดับติดกันที่ผ่านการปรับตำแหน่งแล้ว ส่วนความกว้างของฝ่ามือจะคำนวณจากตำแหน่งร่องนิ้วที่มีลำดับแรกกับลำดับสุดท้ายที่ผ่านการปรับตำแหน่งแล้ว โดยใช้สมการที่ 1 ภาพที่ 4(D) แสดงการหาตำแหน่งกลางนิ้วมือบนเส้นฐานนิ้วมือเพื่อใช้เป็นจุดอ้างอิงในการหาความยาวนิ้ว โดยได้แสดงภาพประกอบเพิ่มเติมในภาพที่ 5(A) ซึ่งจะได้สมการที่ 2 และสมการที่ 3 นำมาหาค่าตำแหน่ง ภาพที่ 4(E) แสดงการหาตำแหน่งกลางนิ้วมือบนเส้นฝ่ามือเพื่อใช้เป็นจุดอ้างอิงในการหาความยาวนิ้ว ได้จากการนำตำแหน่งปลายนิ้วและตำแหน่งกลางนิ้วมือบนเส้นฐานนิ้วมือของแต่ละนิ้วมาคำนวณ โดยได้แสดงภาพประกอบเพิ่มเติมในภาพที่ 5(B) ซึ่งจะได้สมการที่ 6 ที่มาจากการแปลงสมการที่ 4 และสมการที่ 5 นำมาหาค่าตำแหน่ง ภาพที่ 4(F) แสดงการหาความยาวของนิ้วมือ นำจุดทั้งสองที่ได้มาคำนวณหาความยาวของนิ้วดังสมการที่ 1



ภาพที่ 5: (บน) ประกอบสมการที่ 2-3 (ล่าง) ประกอบสมการที่ 4-6

$$X_{center} = \left(\frac{X_{left} + X_{right}}{2} \right) \quad (2)$$

$$Y_{center} = \left(\frac{Y_{left} + Y_{right}}{2} \right) \quad (3)$$

$$\frac{X_{tip} - X_{palm}}{X_{tip} - X_{base}} = \frac{Y_{tip} - Y_{palm}}{Y_{tip} - Y_{base}} \quad (4)$$

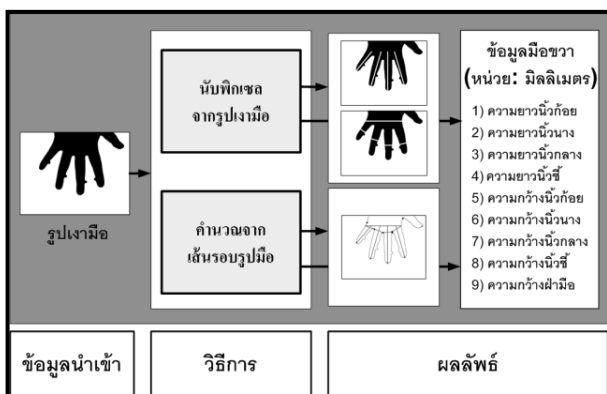
$$\frac{X_2}{X_1} = \frac{Y_2}{Y_1} \quad (5)$$

$$X_2 = \left(\frac{Y_2}{Y_1} \right) \cdot X_1 \quad (6)$$

เมื่อ X คือ พิกัดในแนวแกนนอนในระบบพิกัดฉาก
 Y คือ พิกัดในแนวแกนตั้งในระบบพิกัดฉาก
 left, right, center คือ จุดซ้าย จุดขวา และจุดกลางที่จะหาค่า
 tip, palm, base คือ จุดปลายนิ้ว จุดบนฝ่ามือ และจุดฐานนิ้ว

3. ขั้นตอนวิธีการดำเนินการวิจัย

งานวิจัยนี้จะเน้นไปที่เทคนิคการหาลักษณะเด่นของมือระหว่างวิธีการนับพิกเซลจากรูปเงามือกับวิธีการคำนวณจากเส้นรอบรูปมือ จึงได้ขั้นตอนการทดลองเพื่อเปรียบเทียบหาผลลัพธ์ของทั้งสองวิธี ดังภาพที่ 6 โดยทั้งสองวิธีจะใช้ข้อมูลรูปเงามือเดียวกันเพื่อนำไปหาความยาวนิ้ว, ความกว้างของแต่ละนิ้วและความกว้างฝ่ามือและแสดงผลเป็นภาพ, ข้อมูลมือขวา และเวลาในการประมวลผล



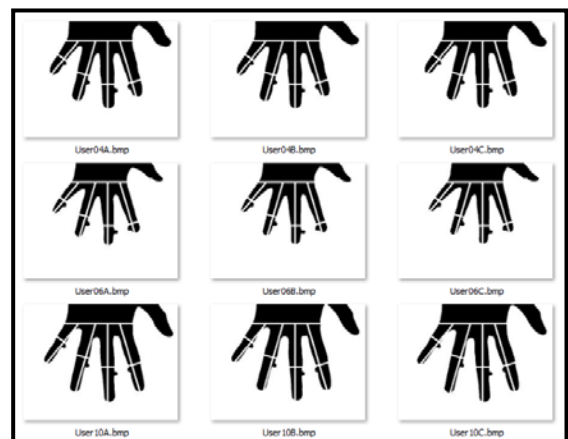
ภาพที่ 6: ขั้นตอนการทดลองเพื่อเปรียบเทียบหาผลลัพธ์

จากขั้นตอนการดำเนินการนำมาเขียนเป็น โปรแกรมที่พัฒนาด้วยภาษา C++ ใช้คอมไพเลอร์ gcc4.4.1 ร่วมกับฟังก์ชันการประมวลภาพของ OpenCV 2.1 โดยกำหนดให้รับข้อมูลจากภาพรูปเงามือขนาด 640x480 พิกเซล และผลลัพธ์ของ

โปรแกรมให้แสดงข้อมูลมือขวามีหน่วยขนาดนิ้วมือเป็นมิลลิเมตร และภาพรูปเงามือพร้อมเส้นแสดงความยาวนิ้วมือและเส้นแสดงความกว้างนิ้วมือเพื่อความสะดวกในการตรวจสอบความถูกต้อง ในการทดสอบนี้ได้นำภาพถ่ายมือที่ได้จากนักศึกษาจำนวน 30 คน คนละ 3 รูป โดยไม่นำข้อมูลนิ้วหัวแม่มือมาพิจารณา จากนั้นนำผลลัพธ์มาวิเคราะห์หาบทสรุปของการทดลอง

4. ผลการดำเนินการ

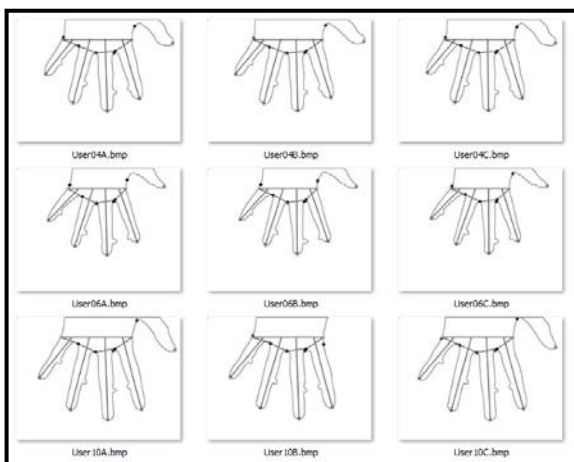
จากขั้นตอนการดำเนินการที่กล่าวมาข้างต้น โปรแกรมได้ทำการประมวลผลบนเครื่องคอมพิวเตอร์ที่ติดตั้งระบบปฏิบัติการ Ubuntu 10.04 มีหน่วยประมวลผลกลาง Intel Pentium Dual ความเร็ว 3 GHz หน่วยความจำหลักขนาด 4 GB ได้ภาพผลลัพธ์จากวิธีการนับพิกเซลจากรูปเงามือ ดังภาพที่ 7 และได้ภาพผลลัพธ์จากวิธีการคำนวณจากเส้นรอบรูปมือนิ่งดังภาพที่ 8 โดยทั้งสองภาพเป็นตัวอย่างภาพผลลัพธ์ของบุคคลทดสอบที่ 4, 6 และ 10 อย่างละสามภาพตามลำดับ และได้ข้อมูลของมือขวาของแต่ละบุคคลจากทั้งสองวิธี นำมาสร้างตารางเปรียบเทียบขนาดของข้อมูลมือขวาของแต่ละบุคคล และตารางเปรียบเทียบความยาวของข้อมูลมือขวา ดังตารางที่ 1 และตารางที่ 2 ตามลำดับ



ภาพที่ 7: ตัวอย่างภาพผลลัพธ์จากวิธีการนับพิกเซลจากรูปเงามือ

จากการสังเกตภาพผลลัพธ์ทั้งหมดด้วยสายตาพบว่า วิธีการคำนวณจากเส้นรอบรูปมือนั้นได้ภาพเส้นแสดงความยาวนิ้วมือที่ดีกว่าโดยเฉพาะในส่วนของเส้นแสดงความยาวนิ้วก้อย เนื่องจากพบว่าภาพผลลัพธ์ในบางบุคคลของวิธีการนับพิกเซล

จากรูปเงามือได้เส้นแสดงความขาวนิ้วก้อยที่ไม่ต่อเนื่อง (เส้นนำร่องขาดบางส่วน) ในส่วนของเส้นแสดงความกว้างนั้นการสังเกตไม่สามารถเปรียบเทียบเส้นแสดงความกว้างของทั้งสองวิธีการงานวิจัยนี้ได้โดยตรง เนื่องจากวิธีการนับพิกเซลจากรูปเงามือนั้นใช้เส้นนำร่องความกว้างนิ้วมือวัดได้ตำแหน่งหมุดกำกับนิ้วมือ แต่วิธีการคำนวณจากเส้นรอบรูปมือใช้เส้นฐานนิ้วมือซึ่งเป็นตำแหน่งที่ต่างกัน แต่เมื่อแยกพิจารณาแล้วไม่พบภาพผลลัพธ์ใดแสดงเส้นความกว้างผิดพลาด ส่วนเส้นแสดงความกว้างฝ่ามือนั้นแสดงผลได้ถูกต้องทั้งหมด ทั้งนี้เนื่องมาจากวิธีการทั้งสองในงานวิจัยนี้ใช้เส้นนี้เป็นจุดอ้างอิงหลัก



ภาพที่ 8: ตัวอย่างภาพผลลัพธ์จากวิธีการคำนวณจากเส้นรอบรูปมือ

จากการวิเคราะห์ตารางเปรียบเทียบขนาดของข้อมูลมือขาวส่วนบุคคล พบว่าข้อมูลขนาดของมือของแต่ละบุคคลทั้งสามตัวอย่าง คือ A, B และ C ซึ่งหมายถึงภาพมือขาวที่ได้จากการบันทึกข้อมูลครั้งที่หนึ่ง, ครั้งที่สอง และครั้งที่สามนั้นสามารถนำมาใช้การยืนยันเอกลักษณ์ได้ เนื่องจากข้อมูลขนาดมือที่ได้ในแต่ละครั้งจากทั้งสองวิธีมีความเที่ยงตรงพอที่จะนำมาใช้ในการจำแนกความแตกต่างของแต่ละบุคคลออกจากกัน ส่วนผลการวิเคราะห์ตารางเปรียบเทียบความยาวของนิ้วมือระหว่างทั้งสองวิธี พบว่าขนาดของความขาวนิ้วมือที่แปลงหน่วยวัดเป็นมิลลิเมตรแล้วจะได้ผลความยาวแทบไม่แตกต่างกันในส่วนของนิ้วกลางและนิ้วชี้ เนื่องจากรูปแบบการวางตำแหน่งของหมุดกำกับนิ้วมือแบบนี้จะบังคับให้นิ้วทั้งสองเคลื่อนตำแหน่งได้น้อยกว่านิ้วอื่น ๆ ในส่วนของขนาดของความขาวนิ้วนางและนิ้วก้อยที่แตกต่างกันมากกว่านิ้วกลางและนิ้วชี้ มาจากผลของวิธีการคำนวณจากเส้นรอบรูปมือได้ทำการปรับตำแหน่งอ้างอิง

บนเส้นฝ่ามือโดยแปรผันตามขนาดความกว้างของนิ้วมือและฝ่ามือ ซึ่งหากมองจากในบางมุมมองการหาลักษณะเด่นด้วยวิธีการคำนวณจากเส้นรอบรูปมือเสมือนการปรับเส้นนำร่องให้ตรงตามขนาดและรูปร่างมือ หรือเส้นนำร่องแบบพลวัต (Dynamic Guide Lines)

ตารางที่ 1: เปรียบเทียบขนาดของข้อมูลมือขาวส่วนบุคคล (บุคคลที่ 4)

บุคคลที่ 04	การวัด	ส่วนหน่ง	หน่วยวัด						
			พิกเซล			มิลลิเมตร			
			ภาพ A	ภาพ B	ภาพ C	ภาพ A	ภาพ B	ภาพ C	Max-min
วิธีนับพิกเซล	ความยาว	นิ้วก้อย	177.95	180.77	172.96	62	63	61	2
		นิ้วนาง	258.56	259.52	256.36	90	91	90	1
		นิ้วกลาง	281.00	280.00	279.00	98	98	98	0
		นิ้วชี้	240.84	237.68	239.00	84	83	84	1
วิธีจากเส้นรอบรูปมือ	ความยาว	นิ้วก้อย	41.22	41.22	39.83	14	14	14	0
		นิ้วนาง	48.13	49.08	49.08	17	17	17	0
		นิ้วกลาง	51.00	52.00	52.00	18	18	18	0
		นิ้วชี้	50.73	48.54	48.54	18	17	17	1
	ฝ่ามือ	274.00	273.00	276.00	96	96	97	1	

ตารางที่ 2: เปรียบเทียบความยาวของนิ้วมือระหว่างทั้งสองวิธี

บุคคลที่	ภาพ	ความยาวนิ้ว (มิลลิเมตร)											
		นิ้วก้อย			นิ้วนาง			นิ้วกลาง			นิ้วชี้		
		วิธีที่ 1	วิธีที่ 2	แตกต่าง	วิธีที่ 1	วิธีที่ 2	แตกต่าง	วิธีที่ 1	วิธีที่ 2	แตกต่าง	วิธีที่ 1	วิธีที่ 2	แตกต่าง
001	A	56	60	4	89	90	1	99	99	0	88	88	0
	B	58	61	3	89	91	2	99	99	0	88	87	1
	C	61	62	1	91	92	1	100	99	1	87	87	0
002	A	60	61	1	92	92	0	99	100	1	88	88	0
	B	61	61	0	92	92	0	99	100	1	87	88	1
	C	62	63	1	92	93	1	99	99	0	86	87	1
003	A	68	72	4	97	99	2	108	108	0	91	91	0
	B	67	72	5	96	98	2	107	108	1	91	91	0
	C	73	77	4	99	101	2	109	109	0	90	90	0
004	A	60	62	2	90	90	0	98	98	0	84	84	0
	B	61	63	2	90	91	1	98	98	0	83	83	0
	C	59	61	2	89	90	1	97	98	1	84	84	0
005	A	59	68	9	93	96	3	105	106	1	93	94	1
	B	62	69	7	94	97	3	107	107	0	94	95	1
	C	64	70	6	96	98	2	108	108	0	92	92	0
006	A	60	57	3	83	82	1	91	91	0	77	76	1
	B	57	56	1	81	81	0	90	90	0	77	76	1
	C	57	54	3	81	79	2	89	89	0	76	76	0
007	A	64	67	3	98	99	1	110	110	0	98	99	1
	B	61	64	3	95	96	1	108	109	1	98	99	1
	C	65	67	2	98	100	2	110	110	0	98	98	0
008	A	62	65	3	95	97	2	107	107	0	91	91	0
	B	62	66	4	94	97	3	106	106	0	89	90	1
	C	62	66	4	95	96	1	106	106	0	90	91	1
009	A	68	69	1	99	101	2	110	110	0	91	89	2
	B	69	70	1	100	102	2	111	111	0	92	90	2
	C	65	67	2	98	99	1	109	109	0	91	91	0
010	A	71	74	3	102	103	1	113	113	0	102	103	1
	B	71	73	2	101	101	0	111	111	0	99	100	1
	C	74	76	2	104	105	1	114	113	1	102	102	0

(หมายเหตุ: แสดงเฉพาะบุคคลที่ 1-10)

หลังจากการประมวลผลด้วยเครื่องคอมพิวเตอร์ข้างต้นก่อนที่จะดำเนินการวัดความเร็วในการประมวลผลบนระบบสมองกลฝังตัว ผู้วิจัยได้ทำการแก้ไขโปรแกรมโดยตัดการประมวลผลด้านกราฟิกออก จากนั้นนำมาคอมไพล์ด้วย GCC Toolchain เพื่อใช้กับระบบปฏิบัติการ Linux for FriendlyARM บน Mini2440 และบน Tiny6410 จากนั้นทำการทดลองซ้ำทุกตัวอย่างโดยบันทึกเวลาในการประมวลผล และหาเวลาที่ใช้ในการประมวลผลโดยเฉลี่ยจากทั้งสองวิธีได้ผลดังตารางที่ 3

ตารางที่ 3: เวลาที่ใช้ในการประมวลผล

	วิธีนับฟิกเซล จากรูปเงามือ	วิธีการคำนวณจาก เส้นรอบรูปมือ
CPU: Intel Pentium Dual Core E5700 Speed: 3 GHz RAM: 4GB	4-5 msec.	5-6 msec.
CPU: ARM11 Samsung S3C6410 Speed: 533 MHz RAM: 256MB	40-44 msec.	68-72 msec.
CPU: ARM9 Samsung S3C2440 Speed: 400 MHz RAM: 64MB	55-60 msec.	100-105 msec.

จากการวิเคราะห์หาปัจจัยที่มีผลกระทบต่อวิธีการหาลักษณะเด่นของมือในระบบที่มีหมุดกำกับนิ้วมือ พบว่าการวางมือให้อยู่ในตำแหน่งที่มีหมุดกำกับอย่างเคร่งครัดส่งผลดีโดยตรงกับความถูกต้องของผลลัพธ์ที่ได้ โดยอ้างอิงจกตารางเปรียบเทียบขนาดของข้อมูลมือขวส่วนบุคคลทั้งหมดมาเลือกเฉพาะตารางที่ให้ส่วนต่างค่าสูงสุดกับค่าต่ำสุดมากผิดปกติ จากนั้นให้ย้อนกลับไปสังเกตภาพเงามือของบุคคลนั้นด้วยสายตาซึ่งพบว่าเกิดจากการวางมือไม่เคร่งครัดในตำแหน่ง ทำให้ภาพเงามือที่ได้เอียงหรือตำแหน่งการวัดต่ำกว่าปกติ

5. สรุปผลการทดลอง

จากการนำเสนอเทคนิคการหาลักษณะเด่นของมือระหว่างวิธีการเปรียบเทียบรูปเงามือด้วยเส้นนำร่องแบบคงที่ กับวิธีการคำนวณจากเส้นรอบรูปมือ ได้ข้อสรุปว่าวิธีการนับฟิกเซลจากรูปเงามือมีข้อดีด้านความเร็วในการประมวลผล ส่วนวิธีการคำนวณจากเส้นรอบรูปมือมีข้อดีที่ในด้านความยืดหยุ่นสูงสามารถรับข้อมูลที่มีขนาดแตกต่างกันได้ดีกว่า และสามารถนำวิธีการนี้มาพัฒนาการหาลักษณะเด่นของมือในระบบวัดโครงสร้างมือแบบไร้หมุดกำกับได้ พร้อมทั้งเป็นแนวทางในการตัดสินใจนำเทคนิคมาประยุกต์ใช้ในระบบวัดโครงสร้างมือด้วยเทคโนโลยีสมองกลฝังตัว

เอกสารอ้างอิง

[1] N. Duta, "A survey of biometric technology based on hand shape," *Pattern Recognition*, vol. 42, pp. 2797-2806, 2009.

[2] A. K. Singh, A. K. Agrawal, and C. B. Pal, "Hand geometry verification system: A review," *International Conference on Ultra Modern Telecommunications & Workshops (ICUMT '09)*, pp. 1-7, 2009.

[3] A. R. a. S. P. Anil K. Jain, "A Prototype Hand Geometry-based Verification System," *International Conference on Audio and Video based Biometric Person Authentication (AVBPA)*, pp. 166-171, 1999.

[4] P. Varchol and D. Levický, "Using of Hand Geometry in Biometric Security Systems", *Radio Engineering Special Issue: Advanced Digital Signal Processing*, Volume 16, 2007.

[5] R. Sanchez-Reillo, C. Sanchez-Avila, and A. Gonzalez-Marcos, "Biometric identification through hand geometry measurements," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 1168-1171, 2000.

[6] M. A. Ferrer, J. Fabregas, M. Faundez, J. B. Alonso, and C. Travieso, "Hand geometry identification system performance," *IEEE International Carnahan Conference on Security Technology (ICCST 2009)*, pp. 167-171, 2009.

[7] M. A. Ferrer, A. Morales, C. M. Travieso, and J. B. Alonso, "Influence of the pegs number and distribution on a biometric device based on hand geometry," *IEEE International Carnahan Conference on Security Technology (ICCST 2008)*, pp. 221-225, 2008.

[8] M. Golfarelli, D. Maio, and D. Malton, "On the error-reject trade-off in biometric verification systems," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, pp. 786-796, 1997.

[9] A. K. Jain and N. Duta, "Deformable matching of hand shapes for user verification," *Proceedings of International Conference on Image Processing, (ICIP 99)*, vol.2, pp. 857-861, 1999.

[10] C. öden, A. Erçil, V. Taylan Yıldöz, H. Körmözötas, and B. Büke, "Hand Recognition Using Implicit Polynomials and Geometric Features," *Audio and Video Based Biometric Person Authentication*. vol. 2091, Springer: Berlin Heidelberg, pp. 336-341, 2001.

[11] G. Tofighi, S. A. Monadjemi, and N. Ghasem-Aghae, "Rapid hand posture recognition using Adaptive Histogram Template of Skin and hand edge contour," *The 6th Machine Vision and Image Processing Conference (MVIP 2010)*, Isfahan, Iran, pp 1-7, 2010.

[12] D. L. Woodard and P. J. Flynn, "Finger surface as a biometric identifier," *Journal of Computer Vision and Image Understanding*, vol. 100, pp. 357-384, 2005.

อุปกรณ์ตรวจสอบช่องทางจราจรแบบเรียลไทม์สำหรับการขับรถทำงาน บนระบบสมองกลฝังตัว Real-Time Lane Detection for Driving using Embedded System

กิตติศักดิ์ ทองยวน(Kittisak Thongyuan)¹ และมหศักดิ์ เกตุมา (Mahasak Ketcham)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ

คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹ktcom04@hotmail.com, ²maoquee@hotmail.com

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อช่วยลดอุบัติเหตุบนท้องถนนโดยใช้วิธีตรวจสอบช่องทางจราจรในขณะที่ขับรถเพื่อช่วยแจ้งเตือนผู้ขับขี่ในขณะที่ผู้ขับขี่ประมาทหรือไม่ชำนาญ โดยใช้หลักการในการประมวลผลภาพ (Image Processing) มาประยุกต์ใช้งานกับระบบสมองกลฝังตัว โดยใช้กล้องในการบันทึกภาพช่องทางจราจรบริเวณด้านหน้าของรถแบบทันทีทันใดและนำภาพที่บันทึกมาทำการการวิเคราะห์โดยใช้หลักการในการประมวลผลภาพโดยใช้เทคนิค Edge Detection และ Hough Transform ซึ่งเป็นหลักการของการประมวลผลภาพและแสดงการแจ้งเตือนให้ผู้ขับขี่ทราบทางจอแสดงผล โดยใช้ระบบสมองกลฝังตัว STM32 สถาปัตยกรรม 32-bits ARM Cortex-M3 และ RapidSTM32 Blockset ซึ่งเป็นชุดกล้องคำสั่งแบบรูปภาพที่สามารถใช้งานร่วมกับ Matlab/Simulink เป็นตัวประมวลผล งานวิจัยนี้สามารถช่วยแจ้งเตือนผู้ขับขี่ ในขณะที่ขับรถออกนอกหรือคร่อมช่องทางจราจรซึ่งช่วยลดอุบัติเหตุที่จะเกิดขึ้นบนท้องถนนได้

คำสำคัญ: การประมวลผลภาพ ระบบสมองกลฝังตัว

Edge Detection Hough Transform

Abstract

This research aims to reduce road accidents by inspecting lanes and to remind careless and unskilled drivers. There are using the principles of image processing to apply with embedded systems. Image processing systems, technic Edge Detection and Hough Transform are worked by installing camera in front of the

car to capture the situation. The captured image will be analyzed immediately. After that, driver will be reminded by showing up in display. Using embedded STM32 32-bits ARM, Cortex-M3 architecture and RapidSTM32 Blockset, a box set image commander can be used with Matlab / Simulink. This research can help alert the driver while driving outside or straddling lanes, which help reduce road accidents.

Keyword: Image Processing, Edge Detection, Embedded systems, Hough Transform

1. บทนำ

ปัจจุบันปริมาณรถบนท้องถนนมีการเพิ่มมากขึ้นและอุบัติเหตุก็เกิดเพิ่มขึ้นตามมาด้วย โดยการเกิดอุบัติเหตุแต่ละครั้งล้วนแล้วแต่ทำให้เกิดการสูญเสียมากมายไม่ว่าจะเป็นการบาดเจ็บ สูญเสียทรัพย์สิน และอาจถึงขั้นเสียชีวิต ซึ่งในการเกิดอุบัติเหตุต่างๆ นั้นก็เกิดจากหลายสาเหตุไม่ว่าจะเป็นการขับรถเร็ว ประมาท ไม่เคารพกฎจราจร เมาแล้วขับ หรือ ขาดความรู้ความชำนาญในการขับรถ ดังนั้นผู้ที่ขับขี่ต้องใช้ความระมัดระวังสูงในการขับรถบนท้องถนน ปัจจุบันการขับรถออกนอกช่องทางจราจรก็เป็นอีกสาเหตุหนึ่ง ที่ทำให้เกิดอุบัติเหตุได้ ในขณะที่ผู้ขับขี่ประมาท อาจจะทำให้รถวิ่งคร่อมช่องทางจราจรหรือวิ่งออกนอกช่องทางจราจรซึ่งทำให้เกิดขวางช่องทางในการแซงของรถคันหลังและอาจข้ามช่องทางจราจรไปอีกช่องทางหนึ่งที่มีรถวิ่งสวนทาง ทำให้เกิดอุบัติเหตุได้ ดังนั้นการตรวจสอบช่องทางจราจรนั้นจึงเป็นประโยชน์ซึ่งทำให้ผู้ขับขี่สามารถรู้ได้ว่าในขณะที่ขับรถอยู่นั้น ขับตรงช่องทางจราจร

หรือไม่ และเพื่อเป็นการแจ้งเตือนผู้ขับขี่ในการขับรถในขณะที่ผู้ขับขี่เดินเลี้ยวไปคร่อมช่องทางจราจรหรือกำลังจะออกนอกช่องทางจราจร ซึ่งจะช่วยให้ผู้ขับขี่ได้รับความปลอดภัยเพิ่มขึ้น และยังช่วยลดอุบัติเหตุและการสูญเสียต่อผู้ขับขี่เองและผู้ใช้รถใช้ถนนคนอื่นๆ รวมไปถึงทรัพย์สินสาธารณะอื่นๆ ได้

ในการตรวจสอบช่องทางจราจรบนถนนในปัจจุบันได้มีผู้คิดค้นและนำเทคโนโลยีมาประยุกต์ใช้งาน โดยใช้หลักการในการประมวลผลภาพหรือ Image Processing ในเวลาจริงซึ่งได้มีการนำวิดีโอที่บันทึกภาพช่องทางจราจรด้านหน้ารถมาทำการวิเคราะห์บนคอมพิวเตอร์โดยออกแบบโปรแกรมเพื่อตรวจสอบช่องทางจราจรบนถนนโดยใช้ Edge Detection [1] และ Hough Transform [2] ซึ่งเป็นฟังก์ชันที่ใช้ใน Matlab เมื่อผ่านกระบวนการประมวลผลแล้วจะแสดงการแจ้งเตือนเมื่อขับรถออกนอกช่องทางจราจรซึ่งจะแจ้งให้ผู้ขับขี่ทราบและยังมีการใช้เทคนิค Hyperbolas [3] ช่วยในการตีเส้นในการระบุขอบของช่องทางจราจร ซึ่งสามารถทำงานได้ดี โดยในส่วนของ การพัฒนาเพื่อนำมาใช้งานจริงได้มี [4] การใช้กล้องเพื่อถ่ายภาพแบบเรียลไทม์และต่อเข้ากับหน่วยประมวลผล Intel Pentium II 933MHz และจอ LCD Touch Screen โดยใช้ K-means Cluster Algorithm เพื่อตรวจสอบหาเส้นช่องทางจราจร และการใช้มีการใช้ FPGA เป็นหน่วยประมวลผลโดยใช้เทคนิค Robert, Sobel และ Prewitt [5] แต่การนำไปใช้งานจริงยังมีขนาดใหญ่และใช้หน่วยประมวลผลที่มีความเร็วต่ำและสิ้นเปลืองพลังงาน ดังนั้นผู้วิจัยจึงมีแนวคิดในการพัฒนาอุปกรณ์ที่สามารถใช้งานได้จริงและเป็นประโยชน์

โดยแนวคิด ของการวิจัยนี้คือการสร้างอุปกรณ์ตรวจสอบช่องทางจราจรเพื่อช่วยในการแจ้งเตือนผู้ขับขี่ไม่ให้ขับรถออกนอกช่องทางจราจร โดยใช้กล้องวิดีโอในการจับภาพแบบเรียลไทม์ และนำภาพที่ได้มาประมวลผลโดยใช้หลักการ Image Processing ซึ่งจะทำการประมวลผลและส่งการบน Embedded System และมีการแจ้งเตือนผู้ขับขี่ในขณะที่ขับรถผ่านทางจอภาพโดยสามารถนำอุปกรณ์ที่พัฒนาขึ้นไปติดตั้งใช้งานได้จริง

2. ทฤษฎีที่เกี่ยวข้อง

2.1 ระบบสมองกลฝังตัว [6]

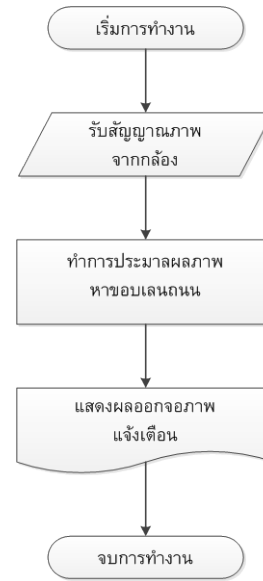
ระบบสมองกลฝังตัว (Embedded System) ก็คือระบบประมวลผล ที่ใช้ชิปหรือไมโครโพรเซสเซอร์ที่ออกแบบมา โดยเฉพาะ เป็นระบบคอมพิวเตอร์ขนาดเล็กที่ฝังไว้ในอุปกรณ์เครื่องใช้ไฟฟ้า และเครื่องเล่นอิเล็กทรอนิกส์ต่างๆ เพื่อเพิ่มความฉลาด ความสามารถให้กับอุปกรณ์เหล่านั้นผ่านซอฟต์แวร์ ซึ่งต่างจากระบบประมวลผลที่เครื่องคอมพิวเตอร์ทั่วไป ระบบฝังตัวถูกนำมาใช้กันอย่างแพร่หลายในยานพาหนะ อุปกรณ์อิเล็กทรอนิกส์ เครื่องใช้ไฟฟ้าในบ้านและสำนักงาน เทคโนโลยีซอฟต์แวร์ เทคโนโลยีฮาร์ดแวร์ เทคโนโลยีเครือข่ายเน็ตเวิร์ก เทคโนโลยีด้านการสื่อสาร เทคโนโลยีเครื่องกลและของเล่นต่าง ๆ คำว่าระบบฝังตัวเกิดจากการที่ระบบนี้เป็นระบบประมวลผลเช่นเดียวกับระบบคอมพิวเตอร์ แต่ว่าระบบนี้จะฝังตัวลงในอุปกรณ์อื่น ๆ ที่ไม่ใช่เครื่องคอมพิวเตอร์ ในปัจจุบันระบบสมองกลฝังตัวได้มีการพัฒนา มากขึ้น โดยในระบบสมองกลฝังตัวจะประกอบไปด้วย ไมโครคอนโทรลเลอร์หรือไมโครโพรเซสเซอร์ อุปกรณ์ที่ใช้ระบบสมองกลฝังตัวที่เห็นได้ชัดเช่นโทรศัพท์มือถือ และในระบบสมองกลฝังตัวยังมีการใส่ระบบปฏิบัติการต่างๆแตกต่างกันไป

2.2 เส้นแบ่งทิศทางจราจร [7]

เส้นแบ่งทิศทางจราจรเป็นเส้นสี่เหลี่ยม อาจเป็นเส้นประหรือเส้นทึบก็ได้ ใช้สำหรับแบ่งทิศทางจราจรของขบวนที่ขับสวนทางกัน โดยรถในแต่ละทิศทางต้องขับด้านซ้ายของเส้นแบ่งทิศทางจราจร ในกรณีเป็นถนน 2 ช่องจราจร เส้นแบ่งทิศทางจราจรจะอยู่ที่ศูนย์กลางทางพอดิ ในกรณีที่ถนนหลายช่องจราจร เส้นแบ่งทิศทางจราจรอาจจะอยู่หรือไม่อยู่ที่ศูนย์กลางก็ได้ ทั้งนี้ขึ้นอยู่กับวิธีจัดการ การจัดจำนวนช่องจราจรให้เหมาะสมกับสภาพการจราจรในแต่ละพื้นที่ เช่น ถนนอาจมีการเพิ่มจำนวนช่องจราจรสำหรับรถวิ่งเข้าไปในบริเวณทางขึ้นเขาที่มีความลาดชันมาก ทำให้มีจำนวนช่องจราจรมากกว่าในทิศทางที่สวนกัน ถนนทางหลวงต้องมีเส้นแบ่งทิศทางจราจร เพื่อความปลอดภัยในการเดินทาง

2.3 การประมวลผลสัญญาณภาพดิจิทัล

การประมวลผลสัญญาณภาพดิจิทัล หรือ Digital Image Processing เป็นการนำภาพเข้าสู่การแปลงข้อมูลภาพให้อยู่ในรูปแบบข้อมูลดิจิทัลที่สามารถนำข้อมูลนี้ผ่านกระบวนการต่างๆเพื่อให้ได้ผลลัพธ์แบบใหม่ที่บ่งบอกถึงลักษณะและคุณสมบัติของภาพ เช่น นำภาพสี RGB แปลงเป็นภาพเฉดขาวดำ (Gray level) การหาขอบภาพการแยกชนิดสี, การดูช่วงค่าความกระจายของสี เป็นต้น กระบวนการต่างๆที่ยกตัวอย่างมานี้ เรียกว่า การกรอง (Filter) การประมวลผลภาพด้วยคอมพิวเตอร์เป็นกระบวนการที่มีเทคนิควิธีในการประมวลผลข้อมูลตัวเลขของภาพที่มีหลากหลายวิธี ซึ่งสามารถนำไปประยุกต์ใช้งานให้เหมาะสมกับข้อมูลภาพที่นำมาประมวลผล โดยปกติข้อมูลภาพจะมีลักษณะเด่นทางด้านรูปร่าง พื้นผิว สี สัน และโครงสร้างต่างๆ ที่แตกต่างกัน ไปขึ้นอยู่กับวัตถุและสภาพแวดล้อมโดยรอบของวัตถุ



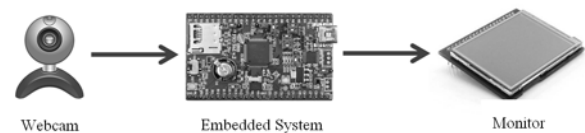
ภาพที่ 1 แสดงขั้นตอนการทำงานของซอฟต์แวร์

3. วิธีที่นำเสนอ

3.1 การออกแบบอุปกรณ์

การทำงานของซอฟต์แวร์จะแสดงให้เห็นดังภาพที่ 1 แสดงขั้นตอน การทำงานของซอฟต์แวร์ซึ่งเป็นขั้นตอนในการออกแบบเพื่อรับข้อมูลเข้ามาโดยข้อมูลจะเป็นสัญญาณภาพที่ได้จากกล้องที่ติดตั้งไว้ที่รถเพื่อจับภาพช่องทางจราจรบนถนนเมื่อรับภาพเข้ามาโปรแกรมจะมีการประมวลผลเพื่อหาขอบและช่องทางจราจร จากนั้นจะทำการวาดเส้นบนช่องทางจราจร เพื่อนำไปแสดงผลทางหน้าจอ โดยจะมีการแจ้งเตือนเมื่อขับรถออกนอกช่องทางจราจร

การออกแบบการทำงานของฮาร์ดแวร์ของอุปกรณ์ตรวจสอบช่องทางจราจรบนถนนแบบเรียลไทม์สำหรับการขับรถจะเป็นการออกแบบดังภาพที่ 2 การต่ออุปกรณ์ โดยจะต่อกับบอร์ดเข้ากับตัว Board Embedded System ซึ่งจะใช้กล้องในการรับภาพเข้ามาประมวลผลและแสดงผลต่างๆ ออกทางจอภาพ โดยตัวกล้องจะถูกติดตั้งไว้ในรถซึ่งเป็นมุมที่สามารถรับภาพช่องทางจราจรบนถนนได้ชัดเจน



ภาพที่ 2 การต่ออุปกรณ์ฮาร์ดแวร์

ขั้นตอนการทำงานของอุปกรณ์ตรวจสอบช่องทางจราจรบนถนนแบบเรียลไทม์ในการขับรถ โดยทำงานบนระบบสมองกลฝังตัว Embedded System

ขั้นตอนที่ 1 การบันทึกภาพ อุปกรณ์ที่ออกแบบจะมีกล้องติดตั้งไว้ที่รถยนต์ไว้คอยบันทึกภาพวิดีโอช่องทางจราจรบนถนน กล้องจะจับภาพและส่งไฟล์เข้ามาที่หน่วยประมวลผล

ขั้นตอนที่ 2 แปลงภาพ หลังจากที่ได้รับไฟล์ ระบบจะทำการวิเคราะห์ไฟล์วิดีโอโดยใช้กระบวนการแยกเฟรม ข้อมูลของไฟล์วิดีโอและถูกเก็บไว้ในอาร์เรย์ซ้อนกันแล้วระบุตัวแปรในการแบ่งเฟรมย่อยในตัวแปรอาร์เรย์ RGB เป็นภาพ YCbCr ตามลำดับเพื่อให้ประมวลผลได้ง่ายขึ้นในขั้นตอนต่อไป

ขั้นตอนที่ 3 การค้นหาเส้นตรง โดยใช้หลักการการค้นหาเส้นตรง Hough Transform เพื่อการค้นหาองค์ประกอบของภาพ ที่มีส่วนเส้นตรงปรากฏในภาพนั้นและเก็บไว้ในอาร์เรย์เป็นการหาเส้นตรงของช่องทางจราจร ในภาพจากฟังก์ชันของสมการ 1

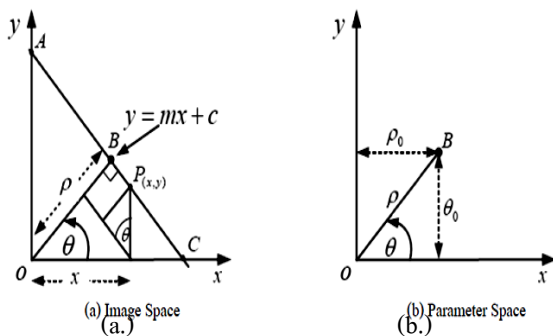
$$f(x, y, \rho_0, \theta_0) = x \cos \theta_0 + y \sin \theta_0 \quad (1)$$

และกำหนดให้จุดของภาพสองมิติมีค่าเท่ากับ $(p-0, \theta-0)$ เมื่อ $p-0$ ระยะที่วัดจากจุดกำเนิดไปตั้งฉากกับเส้นตรง และจุด $\theta-0$ เป็นค่าของมุมระหว่างเวกเตอร์ ρ กับแกน x ดังแสดงในภาพที่ 3 วิธีการ Hough Transform คือจากสมการ 1 มีการกำหนดจุด (x_0, y_0) ของ Image Space ดังนั้นค่าของ ρ จะเท่ากับสมการ 3

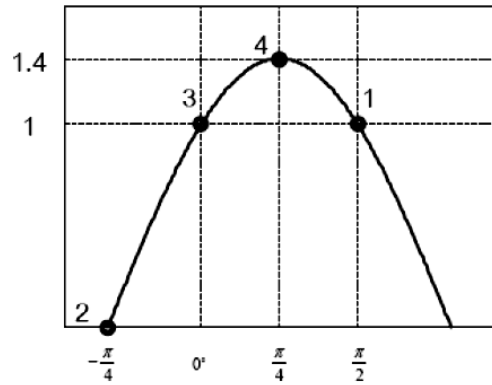
$$\rho = x \cos \theta + y \sin \theta \quad (2)$$

$$\rho = \sqrt{x_0^2 + y_0^2} \left(\frac{x_0}{\sqrt{x_0^2 + y_0^2}} \cos \theta + \frac{y_0}{\sqrt{x_0^2 + y_0^2}} \sin \theta \right) \quad (3)$$

การแปลงรูปแบบของ HT พบว่า $\rho = x \cos \theta + y \sin \theta$ ของจุด (x_0, y_0) ใน Image Space เป็นการแปลงรูปแบบไปสู่เส้นโค้งรูปซายด์ (Sinusoidal Curve) ใน Parameter Space ทิศทางทวนเข็มนาฬิกาและจุด $(p-0, \theta-0)$ ของรูปเส้นโค้งรูปซายด์นี้ แสดงให้เห็นแทนเส้นตรงที่ลากผ่านจุด (x_0, y_0) ใน Image Space เมื่อทดลองแทนค่า θ และ ρ ตามภาพที่ 3 ดังนั้นผลที่ได้จะเป็นการทำ Hough Transform ระหว่างค่าของเส้นตรงใน Image Space ไปสู่ค่าของจุดใน Hough Transform ดังแสดงในภาพที่ 4

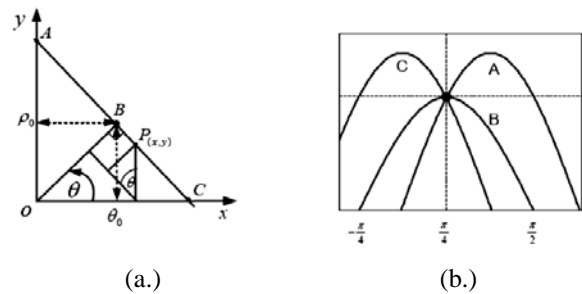


ภาพที่ 3 การนับจำนวนเส้นตรงของการเปลี่ยนแปลง Hough Transform



ภาพที่ 4 ตัวอย่างการทำ Hough Transform

เมื่อทั้ง 4 จุดมีที่เกิดขึ้นใน Hough Space ค่าของจุด 2 3 4 และ 1 ในภาพที่ 5 (b) จะมีค่าตรงกันในแต่ละเส้นของในภาพภาพที่ 5 (a) Image Space ลักษณะของภาพแบบการแปลงบทกลับของ Hough Transform ค่าของจุด $(p-0, \theta-0)$ ใน Parameter Space สามารถอธิบายได้ด้วยการแปลงรูปแบบจาก Spatial Domain ไปสู่การแทนค่าของเส้นตรง



ภาพที่ 5 การแปลงค่า Hough Transform เมื่อมีการเปลี่ยนแปลงจุดในเส้นตรง

จากภาพภาพที่ 5(a.) ตัวอย่างของ 3 จุดของตำแหน่ง A(0,2) B(1,1) และ C(2,0) ของทุกจุดบนเส้นตรง จุดดังกล่าวนี้เป็นจุดที่ตรงใน Hough Space ในความเป็นจริงแล้วจุดเป็นค่าหนึ่งของสมการเส้นตรงเส้นที่สองและสามและมีค่าสอดคล้องกับเส้นโค้งที่รวมอยู่ในจุดของค่าเวกเตอร์ ρ เท่ากับ 1.414 และค่าของมุม θ เท่ากับ 0.79 เมื่อนำค่าของ (ρ, θ) แทนค่าในสมการ 4 ได้ดังนี้

$$1.414 = x \cos(0.79) + y \sin(0.79) \quad (4)$$

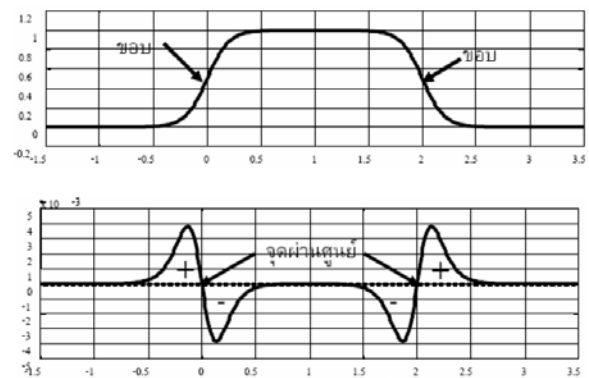
เมื่อได้ทำการ Hough Transform แล้วจุดที่ P(x,y) ที่ปรากฏอยู่ในภาพจะมีเส้นตรงจำนวนมากมาที่ลากผ่านได้ ดังนั้นวิธีการของ Hough Transform คือการนับค่าว่าจุด P(x,y) ดังกล่าวมีจำนวนเส้นของเส้นตรงลากผ่านจุดนี้จำนวนเท่าไร และถ้าพิจารณาใน Hough Space ก็จะมองเห็นเป็นจุดที่มีค่าของเส้นโค้งรูปซายด์ตัดผ่าน แสดงว่าจุดนั้นคือจุดเด่นของเส้นตรงที่ผ่านจุด P(x,y) มีค่ามากที่สุด จุดดังกล่าวมีค่าระยะห่างจากจุดกำเนิดมากที่สุดก็ต่อเมื่อเวกเตอร์ ρ ตั้งฉากกับเส้นตรงที่ผ่านจุด P(x,y) ซึ่งหลังจากประมวลผลจะได้ภาพดังตัวอย่างภาพที่ 6 เมื่อเสร็จสิ้นการหารทุกเฟรมระบบจะเก็บเฟรมในเฟรมบัฟเฟอร์ซึ่งจะปรากฏในโมดูลการจัดเก็บข้อมูลสำหรับการวิเคราะห์ในขั้นตอนต่อไป



ภาพที่ 6 เส้นที่ได้จาก Hough Transform

ขั้นตอนที่ 4 การหาขอบของช่องทางจราจรบนถนน โดยใช้หลักการ Edge Detection คือการตรวจสอบว่าเส้นขอบลากผ่านหรือใกล้กับจุดใด โดยวัดจากการเปลี่ยนแปลงของตำแหน่งที่ใกล้เคียงกับจุดดังกล่าว ซึ่งวิธีการหาขอบนั้นได้ใช้วิธี Laplacian Method คือการค้นหาโดยใช้อนุพันธ์อันดับสองเปรียบได้กับการหาอนุพันธ์อันดับสองของภาพเพื่อให้ได้ขอบของภาพ โดยภาพที่ผ่านการหาอนุพันธ์อันดับสอง (Laplacian Operator($\nabla^2 P$)) บริเวณที่เป็นส่วนขอบจะเด่นชัดขึ้น การประมาณค่าของการหาอนุพันธ์อันดับสองโดยใช้ Mask ของสมการ 6 และ 7 ทำให้ตำแหน่งของบริเวณขอบของวัตถุในภาพคือค่าจุดผ่านศูนย์ (Zero Crossing) ของการหาอนุพันธ์อันดับสอง สามารถหาได้จากสมการ 5 การค้นหาขอบภาพด้วยอนุพันธ์อันดับสอง เมื่อทำ $\nabla^2 P$ บริเวณขอบจะมีตำแหน่ง

เดียวกับค่าจุดผ่านศูนย์ของค่าจาก $\nabla^2 P$ พิจารณาได้จากตำแหน่งที่พิกเซลเปลี่ยนแปลงจากค่าที่เป็นบวกเป็นค่าที่เป็นลบหรือจากค่าที่เป็นลบไปเป็นค่าที่เป็นบวก ดังภาพที่ 7 การหาขอบโดยใช้อนุพันธ์อันดับสอง ไม่สนใจทิศทางของภาพในแนวแกน x และแกน y กำหนดจุดที่ค่า y เป็นจุดผ่านศูนย์ วิธีนี้ใช้เวลาในการค้นหาขอบโดยใช้อนุพันธ์อันดับหนึ่ง กล่าวคือเราสามารถตรวจจับความไม่ต่อเนื่องของพิกเซลในโดเมนรูปภาพได้โดยใช้อนุพันธ์ของภาพนั่นเอง เนื่องจากการจัดเก็บรูปภาพเป็นการจัดเก็บในรูปแบบพิกเซล ฉะนั้นการที่จะจำแนกวัตถุ 2 วัตถุออกจากกันหรือแยกวัตถุออกจากพื้นหลังจึงต้องอาศัยขอบ (Edge) ของวัตถุ การหาขอบของวัตถุสามารถทำได้โดยการดูความแตกต่างของสีของพิกเซลใกล้เคียง ถ้าสีแตกต่างกัน แสดงว่ามีขอบอยู่ระหว่างพิกเซลนั้น



ภาพที่ 7 การหาขอบโดยใช้อนุพันธ์อันดับสอง

$$\nabla^2 P = \frac{\partial^2 P}{\partial x^2} + \frac{\partial^2 P}{\partial y^2} \quad (5)$$

$$\text{Mask}(E_x) = \begin{bmatrix} Z_{x1} & Z_{x2} & Z_{x3} \\ Z_{x4} & Z_{x5} & Z_{x6} \\ Z_{x7} & Z_{x8} & Z_{x9} \end{bmatrix} = E_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad (6)$$

$$\text{Mask}(E_y) = \begin{bmatrix} Z_{y1} & Z_{y2} & Z_{y3} \\ Z_{y4} & Z_{y5} & Z_{y6} \\ Z_{y7} & Z_{y8} & Z_{y9} \end{bmatrix} = E_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad (7)$$



ภาพที่ 8 การหาขอบโดย Edge Detection

ขั้นตอนที่ 5 แสดงผล การแสดงผลจะแสดงบนหน้าจอให้เห็นช่องทางจราจรในขณะที่ขับรถรวมทั้งดีเส้นช่องทางจราจรที่ตรวจจับได้ทั้งซ้ายและขวาดังภาพที่ 9 และภาพที่ 10 โดยมีการแจ้งเตือนเมื่อรถขับชิดช่องทางจราจรทางซ้ายหรือขวาเกินไปหรือกำลังจะออกนอกช่องทางจราจรที่แสดงบนจอภาพ



ภาพที่ 9 การแสดงผลเมื่อรถชิดช่องทางจราจรทางขวาเกินไป



ภาพที่ 10 การแสดงผลเมื่อรถชิดช่องทางจราจรทางซ้ายเกินไป

โดยระบบจะทำการวิเคราะห์ว่ารถอยู่ในตำแหน่งกลางของช่องทางจราจรหรือไม่และทำการวัดระยะจากขอบทั้งสองข้างหากรถเอียงออกนอกช่องทางจราจรระบบจะแจ้งเตือนผ่านจอภาพ

ขั้นตอนที่ 6 ระบบทำงานวนไปเรื่อยๆ และรับภาพเข้ามาตามขั้นตอนที่ 1 จนถึงขั้นตอนที่ 5 และวนซ้ำไปตลอด

4. สรุปผล

อุปกรณ์ตรวจสอบช่องทางจราจรบนถนนแบบเรียลไทม์สำหรับการขับรถ โดยทำงานบน Embedded System สามารถช่วยแจ้งเตือนให้ผู้ขับขี่ให้ระมัดระวังในการขับโดยสามารถตรวจสอบได้ว่ารถขับออกนอกช่องทางจราจรหรือไม่และมีการตรวจสอบช่องทางจราจรแบบเรียลไทม์ทำให้สามารถแจ้งเตือนได้ในขณะที่รถขับคร่อมช่องทางจราจรหรือกำลังจะออกนอกช่องทางจราจร ซึ่งช่วยเพิ่มความระมัดระวังให้ผู้ขับขี่และช่วยลดอุบัติเหตุบนท้องถนนได้

เอกสารอ้างอิง

- [1] G.Thittaporn, and K.Mahasak, and S.Sasipa, "Real-Time Lane Detection for Driving System Using Image Procissing Base on Edge Detection and Hough Transform," *The Third International Conference on Digital Information and Communication Technology and its Applications*, 2013.
- [2] P.Thitapa, K.Mahasak, "A Lane Detection for the Driving System Based on the Histogram Shapes," *International Conference on Systems and Electronic Engineering (ICSEE'2012)*, Phuket, pp.43-47, December 18-19, 2012.
- [3] A. Abdulhakam, O. Khalifa, I. Rafiqul, K. Sheroz, "Read Time Lane Detection for Autonomous Vehicles," *Proceedings of the International Conference on Computer and Communication Engineering 2008*. pp. 82-88, 2008.
- [4] M.Jer Jeng, P.Chih Hsueh, C.Wei Yeh, P.Yung Hsiao1, C.Cheng, and L.Be Chang, "Real time mobile lane detection system on PXA255 embedded system," *International Journal of Circuits, Systems and Signal Processing*, pp. 177-181,2007.
- [5] K.Warut "Hardware / Software Co Design for Line Detection Algorithm on FPGA," *Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, (ECTI-CON2009)*, 2009.
- [6] ผศ. สุพนธ์ คุ่มนาค "ระบบสมองกลฝังตัว" ภาควิชาวิศวกรรมคอมพิวเตอร์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี
- [7] กรมทางหลวง "คู่มือและมาตรฐานเครื่องหมายจราจรบนพื้นทางและเครื่องหมายนำทาง" กระทรวงคมนาคม 2554.

การออกแบบโครงสร้างแป้นพิมพ์เสมือนด้วยการคำนวณเชิงวิวัฒนาการ

Design Virtual Keyboard Layout Using Evolutionary Computation

วีรวรรณ จันทนะทรัพย์ (Veerawan Janthanasub)¹ พยุง มีสัจ (Phayung Meesad)² และเมธา ศิริกุล (Metha Sirigool)³

¹คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร

²คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

³คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
ratree.j@rmutp.ac.th, pym@kmutnb.ac.th, metha@mail.rmutt.ac.th

บทคัดย่อ

บทความนี้นำเสนอการใช้อุปกรณ์คอมพิวเตอร์ในการออกแบบตำแหน่งอักษรภาษาไทยบนแป้นพิมพ์เสมือน โดยใช้ค่าความถี่ของอักษร และค่าความถี่ของคู่อักษร คำนวณหาตำแหน่งที่เหมาะสมด้วยวิธีคำนวณเชิงวิวัฒนาการ โดยมุ่งหวังเพื่อเพิ่มความเร็วในการพิมพ์ข้อความ ออกแบบและทดสอบด้วยข้อมูลขนาด 5 ล้านคำ จากโครงการวัดเปรียบเทียบสมรรถนะเพื่อพัฒนามาตรฐานการประมวลผลภาษาไทย ผลการวิจัยแสดงให้เห็นว่าโครงสร้างแป้นพิมพ์เสมือนภาษาไทยที่ออกแบบใหม่ให้ประสิทธิภาพความเร็วในการพิมพ์เฉลี่ย 41 คำต่ออนาที และจากการวิเคราะห์โครงสร้างแป้นพิมพ์ใหม่สามารถพิมพ์ได้รวดเร็วกว่าการใช้แบบพิมพ์ที่มีอยู่ในปัจจุบัน

คำสำคัญ: การคำนวณเชิงวิวัฒนาการ แป้นพิมพ์เสมือน และโครงสร้างแป้นพิมพ์

Abstract

This paper presents a computer aided design technique for the virtual keyboard layout. The frequency of characters and bigram together with evolutionary computation used to determine the layout. The proposed keyboard layout is augmented with word prediction list derived from Benchmark for Enhancing the Standard for Thai Language corpus to speed up text entry (41 wpm). Performance analysis of new virtual keyboard layout performs better than traditional layout.

Keyword: Evolutionary Computation, Virtual Keyboard, Keyboard Layout

1. บทนำ

การปฏิสัมพันธ์ระหว่างมนุษย์กับอุปกรณ์คอมพิวเตอร์ในการพิมพ์อักษร (alphabet) หรือข้อความ (word) เปลี่ยนไปเนื่องจากปัจจุบันอุปกรณ์คอมพิวเตอร์ส่วนใหญ่มีหน้าจอเป็นระบบสัมผัส จากเดิมที่สั่งการให้คอมพิวเตอร์ทำงานผ่านแป้นพิมพ์ หรือคีย์บอร์ด (computer keyboard) เปลี่ยนไปเป็นสั่งการโดยการสัมผัสนิ้วบนจอภาพแทน การพิมพ์ผ่านจอภาพระบบสัมผัสเป็นพิมพ์ถูกโปรแกรมให้อยู่ในรูปของซอฟต์แวร์ (software) ซึ่งเรียกว่า แป้นพิมพ์เสมือน (virtual keyboard) ผู้พิมพ์ใช้นิ้วเพียงนิ้วเดียวสัมผัสบนปุ่มแป้นพิมพ์ที่ถูกโปรแกรมไว้ แต่สำหรับโครงสร้างแป้นพิมพ์คอมพิวเตอร์ที่ใช้ในงานในปัจจุบัน ตำแหน่งของแป้นพิมพ์ถูกออกแบบและพัฒนามาจากเครื่องพิมพ์ดีดในอดีต (typewriter) ซึ่งตำแหน่งของแป้นอักษรบนเครื่องพิมพ์ดีดถูกออกแบบมาสำหรับการพิมพ์ระบบสัมผัสสปีดนิ้ว และเมื่อผู้ใช้พิมพ์ข้อความแล้วก้านพิมพ์ไม่เกิดการติดขัด

จะเห็นได้ว่าเป็นพิมพ์คอมพิวเตอร์ และเป็นพิมพ์เสมือนมีรูปแบบการปฏิสัมพันธ์ในการพิมพ์อักษรที่แตกต่างกัน ดังนั้นโครงสร้างตำแหน่งของแป้นอักษรบนแป้นพิมพ์คอมพิวเตอร์จึงอาจไม่เหมาะสมกับการพิมพ์บนจอภาพระบบสัมผัสเนื่องจากตำแหน่งของแป้นพิมพ์ที่เหมาะสมจะส่งผลทำให้การพิมพ์มีประสิทธิภาพสูง สำหรับงานวิจัยนี้ศึกษาวิธีการแก้ปัญหาการใช้งานแป้นพิมพ์เสมือนภาษาไทยให้มีประสิทธิภาพดียิ่งขึ้น โดยการวิเคราะห์ค่าความถี่ของอักษร และความถี่ของคู่อักษร คำนวณหาตำแหน่งที่เหมาะสมของแป้นพิมพ์แต่ละตัว ผลลัพธ์ที่ได้ คือ โครงสร้างแป้นพิมพ์

เสมือนภาษาไทยใหม่ที่ทำให้ประสิทธิภาพความเร็วในการพิมพ์ที่ดีกว่าเดิม ในการเรียบเรียงบทความงานวิจัยฉบับนี้ ผู้วิจัยได้แบ่งหัวข้อบทความเป็นดังนี้ ปัญหาการออกแบบโครงสร้างเป็นพิมพ์เสมือน วิธีการดำเนินงานวิจัย ผลการวิจัย และลำดับสุดท้ายเป็นการสรุปผลงานวิจัยและอภิปรายผล

2. ปัญหาการออกแบบโครงสร้างเป็นพิมพ์เสมือน (Problem of Design Virtual Keyboard Layout)

จากการศึกษาหลักการ ทฤษฎี และงานวิจัยที่เกี่ยวข้องพบว่า กระบวนการออกแบบโครงสร้างเป็นพิมพ์เสมือนประกอบด้วย 4 ขั้นตอนหลักคือ (1) เตรียมคลังข้อมูล (2) วิเคราะห์เชิงความถี่ของข้อมูลอักษร คู่อักษร และวิเคราะห์คำที่ใช้งานบ่อยในแต่ละภาษา (3) ออกแบบโครงสร้างของเป็นพิมพ์ และ (4) คำนวณหาตำแหน่งเป็นพิมพ์ที่เหมาะสม

2.1 การเตรียมคลังข้อมูล (preparation data of corpus) กระบวนการเตรียมคลังข้อมูล คือ การเตรียมข้อมูลเอกสารข้อความ (text) ของแต่ละภาษา ซึ่งต้องมีจำนวนคำที่มากพอและมาจากหลายหมวดหมู่ของเอกสาร อีกทั้งมีความเป็นปัจจุบันอีกด้วย กระบวนการเริ่มต้นจากรวบรวมเอกสารข้อความ แล้วป้อนเข้าสู่กระบวนการตัดคำ (word segmentation) สำหรับภาษาไทยปัจจุบันมีซอฟต์แวร์ที่ใช้เป็นเครื่องมือการตัดคำข้อความภาษาไทย (Thai word segmentation tool) ที่นิยมใช้ในปัจจุบันคือ โปรแกรมตัดคำและกำกับหน้าที่คำภาษาไทยอัตโนมัติ (Smart Word Analysis for Thai, SWATH) โปรแกรมตัดคำสำหรับข้อความภาษาไทย (Thai Lexeme Tokenizer, LexTo) หรือ เล็กซ์โต และโปรแกรมตัดคำสำหรับข้อความภาษาไทย (Thai Lexeme Analyzer, TLexs) หรือ ทีเล็กซ์ ซึ่งทั้งสามโปรแกรมพัฒนาภายใต้การควบคุมของศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ และผลลัพธ์ที่ได้จากการตัดคำจะถูกส่งไปยังกระบวนการวิเคราะห์ความถี่คำและอักษรต่อไป

2.2 การวิเคราะห์ความถี่คำและอักษร (word and letter frequency analysis) ในการคำนวณหาตำแหน่งของเป็นพิมพ์ที่เหมาะสม ต้องอาศัยความสัมพันธ์ของอักษรที่ประกอบกันเป็นคำในแต่ละภาษา ดังนั้นข้อมูลค่าความถี่ของอักษรแต่ละตัวและค่าความถี่ของคู่อักษร จึงถูกนำมาใช้ในการคำนวณ

สำหรับเทคนิคที่ใช้วิเคราะห์ความถี่ของอักษรที่นิยมใช้กันมากที่สุดคือ แบบจำลองเอ็น-แกรม (n-Gram) เพราะเป็นวิธีการที่เรียบง่าย มีประสิทธิภาพสูง และเหมาะสมสำหรับวิเคราะห์ภาษา สามารถใช้ระบุภาษาได้ดีกว่าวิธีอื่น ๆ หลักการทำงานคือ นำข้อมูลจากคลังข้อมูลที่จัดเตรียมไว้ มาดำเนินการคำนวณหาค่าความน่าจะเป็นของชุดอักษร (letter sequence) ที่เกิดขึ้นร่วมกันเป็นคำ (word) แกรม (Gram) คือ หน่วยที่ใช้สร้างแบบจำลอง ซึ่งแกรมมีได้หลายขนาดแล้วแต่จะกำหนดตั้งแต่ 1 จนถึงเอ็น (n) สำหรับปัญหาการออกแบบแผนผังเป็นพิมพ์เสมือน จะใช้ 1 แกรม (probability monogram) และ 2 แกรม (probability bigram/diagram) เพื่อคำนวณหาค่าความน่าจะเป็นของอักษรแต่ละตัวและคู่อักษร

2.3 การออกแบบโครงสร้างของเป็นพิมพ์เสมือน เป็นการออกแบบจำนวนแถว (row) และจำนวนคอลัมน์ (column) รูปทรงของเป็นพิมพ์ (shape) ขนาดของเป็นพิมพ์ (size) และจำนวนหน้าของเป็นพิมพ์ (pages) สำหรับโครงสร้างเป็นพิมพ์คอมพิวเตอร์มาตรฐานภาษาอังกฤษใช้แบบพิมพ์ควอร์ตี (QWERTY) และดีโวเร็ค (Dvorak) [1] และมีนักวิจัยได้ศึกษาและออกแบบเป็นพิมพ์เหมือนภาษาอังกฤษโดยมีวัตถุประสงค์หลักเพื่อเพิ่มประสิทธิภาพของเป็นพิมพ์ตัวอย่างเช่น แบบพิมพ์ FITALY [2] แบบพิมพ์ OPT I และ OPT II [3] เป็นต้น

สำหรับโครงสร้างเป็นพิมพ์ภาษาไทยมีความยุ่งยากกว่าภาษาอังกฤษ เนื่องจากภาษาไทยมีความซับซ้อนมากเมื่อเทียบกับเป็นภาษาอังกฤษ ภาษาไทยมีรากฐานมาจากภาษาไทกะได (Tai-Kadai Language Family) อักษรไทยเขียนในแนวเชิงเส้นเรียงลำดับจากซ้ายไปขวา และจากบนลงล่าง ไม่มีสัญลักษณ์ในการแบ่งคำ อักษรในภาษาไทยประกอบด้วย 87 อักษร [4] สามารถจำแนกอักษรในภาษาไทยออกเป็นกลุ่มพยัญชนะ 44 ตัว กลุ่มสระ 18 ตัว วรรณยุกต์ 4 ตัว กลุ่มกำกับเสียง จำนวน 5 ตัว กลุ่มตัวเลข 10 ตัว และกลุ่มอักษรพิเศษจำนวน 6 ตัว ดังนั้นการพิมพ์ข้อความภาษาไทยจึงใช้เวลาในการพิมพ์นานกว่าข้อความภาษาอังกฤษ โดยเป็นพิมพ์ภาษาไทยที่นิยมใช้มีอยู่ด้วยกัน 2 แบบ คือ แบบพิมพ์เกษมณี (Ketmanee Layout) และแบบพิมพ์ปัตตะโชติ (Pattajoti Layout) แต่ที่นิยมใช้มากที่สุดคือ แบบพิมพ์เกษมณี ในกระบวนการ

ผลิตของผลิตภัณฑ์อุตสาหกรรมเป็นพิมพ์คอมพิวเตอร์ ถูกกำหนดมาตรฐานการผลิตชื่อ มอก.820 (TIS-820) โดยสถาบันมาตรฐานผลิตภัณฑ์อุตสาหกรรม (Thai Industrial Standard Institute, TISI) และได้กำหนดให้เป็นพิมพ์เกษมณี เป็นพิมพ์มาตรฐานของภาษาไทย [4]

แบบพิมพ์เกษมณี ออกแบบโดย สุวรรณประเสริฐ (กิมเฮง) เกษมณี ใช้เวลาออกแบบนานกว่า 7 ปี จากการรวบรวมเอกสาร ตำรา หนังสือในสมัยนั้น สำหรับเป็นพิมพ์ปัตตะโชติ ถูกวิจัยและออกแบบโดย สกฤษฎ์ ปัตตะโชติ จากผลการวิจัยพบว่าแบบพิมพ์ปัตตะโชติสามารถพิมพ์ได้เร็วกว่าแบบพิมพ์เกษมณีถึงร้อยละ 25.8 แต่ไม่ได้รับความนิยมเนื่องจากผู้ใช้งานมีความเคยชินกับแบบพิมพ์เกษมณี และการปรับเปลี่ยนโครงสร้างของแบบพิมพ์ของเครื่องพิมพ์ดีดมีค่าใช้จ่ายสูงในสมัยนั้น อย่างไรก็ตามแบบพิมพ์เกษมณีออกแบบมาสำหรับเครื่องพิมพ์ดีดในระบบพิมพ์สัมผัสสลิปนิ้วจึงอาจไม่มีความเหมาะสมกับการพิมพ์สัมผัสเพียงนิ้วเดียวบนจอภาพระบบสัมผัสของเป็นพิมพ์เสมือน [5]

จากปัญหาข้างต้นมีนักวิจัยศึกษาและออกแบบโครงสร้างเป็นพิมพ์ภาษาไทยเสมือนใหม่ เช่น งานวิจัยของ ไกรกรรชิต และเกริก [6] ออกแบบโครงสร้างเป็นพิมพ์เสมือนรูปทรงสี่เหลี่ยมผืนผ้า (rectangle layout) ขนาด 9 แถว 11 คอลัมน์ โดยใช้ขั้นตอนวิธีพันธุกรรม (Genetic Algorithm, GA) กำหนดฟังก์ชันวัดค่าความแข็งแรง (fitness function) คือ กฎของฟิต (Fitts' Law) ใช้คำนวณเวลาในการพิมพ์ผ่านเป็นพิมพ์เสมือน โดยใช้ตัวอย่างคำภาษาไทยจากการสุ่มเอกสารบนเว็บไซต์ที่เผยแพร่ทั่วไปจำนวน 200,000 อักษรผลลัพธ์ที่ได้ คือ โครงสร้างเป็นพิมพ์เสมือนใหม่ที่ใช้เวลาในการพิมพ์น้อยกว่าโครงสร้างเป็นพิมพ์แบบเกษมณีและปัตตะโชติ ต่อมา จิรภาต และเกริก [7] นำเสนอการปรับปรุงการใช้งานเป็นพิมพ์เสมือนภาษาไทยบนอุปกรณ์หน้าจอสัมผัสให้เหมาะสมกับการใช้งานบนเครือข่ายทางสังคมออนไลน์ (social network) โดยออกแบบเป็นพิมพ์รูปทรงสี่เหลี่ยมจัตุรัส (square layout) ขนาด 10 แถว 10 คอลัมน์ โครงสร้างเป็นพิมพ์เสมือนประกอบด้วยพยัญชนะ สระ วรรณยุกต์ภาษาไทย ตัวเลขอารบิกศูนย์ถึงเก้า (0-9) และเครื่องหมายวรรคตอน ทดสอบกับข้อมูลที่รวบรวมจากเว็บไซต์เครือข่าย

ทางสังคม คือ Facebook, Twitter และ Hi5 ขนาดข้อมูล 30,000 คำ จากนั้นดำเนินการวัดประสิทธิภาพของเป็นพิมพ์เสมือน โดยใช้กฎของฟิต (Fitts' law) ในการคำนวณเวลาในการพิมพ์ ผลการวิจัยพบว่าโครงสร้างเป็นพิมพ์ใหม่ที่ออกแบบขึ้นให้ประสิทธิภาพความเร็วในการพิมพ์ดีกว่าเป็นพิมพ์เดิมอื่น ๆ

2.4 การคำนวณหาตำแหน่งเป็นพิมพ์ที่เหมาะสมเพื่อเพิ่มประสิทธิภาพในการพิมพ์ให้เร็วขึ้น จัดอยู่ในกลุ่มปัญหาที่มีความซับซ้อนที่ยากต่อการแก้ไขหากค่าตอบ (Nondeterministic Polynomial time, NP-Complete) ใดๆก็ตามปัญหาในกลุ่มที่มีความซับซ้อนและหาค่าตอบยากมักถูกแก้ปัญหาโดยวิธีการคำนวณเชิงวิวัฒนาการ (evolutionary computation) ซึ่งเป็นหนึ่งในศาสตร์ของปัญญาประดิษฐ์ (Artificial Intelligence, AI) โดยขั้นตอนวิธีการคำนวณเชิงวิวัฒนาการนั้นใช้กระบวนการที่ได้รับแรงบันดาลใจมาจากการวิวัฒนาการทางชีววิทยา [8]

วิธีการคำนวณเชิงวิวัฒนาการ ประกอบด้วยกระบวนการหลัก 6 กระบวนการ สามารถอธิบายได้ดังนี้ เริ่มต้นจากสร้างโครโมโซมประชากรเริ่มต้น จากนั้นนำโครโมโซมประชากรแต่ละตัวประเมินค่า นำคำตอบที่ได้ประเมินด้วยฟังก์ชันวัตถุประสงค์ (objective function) เพื่อคัดเลือก (selection) เป็นประชากรต้นกำเนิดสายพันธุ์ต่อไป ดำเนินการแลกเปลี่ยนยีนส์ (recombination) หรือกลายพันธุ์ (mutation) จนได้โครโมโซมลูกหลาน และนำไปแทนที่โครโมโซมประชากรต้นกำเนิด (replacement) ดำเนินการวนซ้ำจนค่าที่เหมาะสม (fitness value) ที่คำนวณได้ไม่มีการเปลี่ยนแปลง จะเห็นว่าโครโมโซมในแต่ละรุ่นถูกปรับปรุงด้วยตัวดำเนินการทางพันธุกรรม ทำให้ค่าความเหมาะสมจะดีขึ้นเรื่อยๆ ในแต่ละรอบ จึงเรียกวิธีการนี้ว่าการคำนวณเชิงวิวัฒนาการนั่นเอง หัวใจสำคัญของการแก้ปัญหาเพื่อหาคำตอบที่เหมาะสมคือ ฟังก์ชันวัตถุประสงค์ (objective function) สำหรับปัญหาการคำนวณตำแหน่งเป็นพิมพ์ที่เหมาะสม ฟังก์ชันวัตถุประสงค์ที่นำมาใช้คือ สมการกฎของฟิต (Fitts' law equation) ถูกนำเสนอโดย Paul Fitts ในปี ค.ศ. 1954 [9] ภายใต้แนวคิดจำลองพฤติกรรมตอบสนองของร่างกายมนุษย์ (human psychomotor behavior) ในการคาดการณ์เวลาการเคลื่อนที่ไปยังวัตถุเป้าหมาย สามารถประยุกต์กฎของฟิตกับการเคลื่อนที่

ของนิ้วมือในการพิมพ์บนแป้นพิมพ์เสมือน สามารถนิยามกฎของฟิต ดังสมการที่ 1

$$MT = a + b \log_2 \left(\frac{D_{ij}}{W_j} + 1 \right) \quad (1)$$

โดยที่ a, b คือ ค่าสัมประสิทธิ์ที่ได้จากการถดถอย
 i, j คือ ตำแหน่งเป็นอักษรต้นทางและปลายทาง
 D_{ij} คือ ระยะห่างระหว่างเป็นอักษร
 W_j คือ ขนาดของแป้นอักษร

สำหรับการพิมพ์ข้อความ โดยใช้ข้อมูลเชิงความถี่ของตัวอักษร และคู่อักษรในการพิมพ์ข้อความ สามารถพยากรณ์เวลาเฉลี่ยในการพิมพ์ (\bar{t}) ดังนิยามในสมการที่ 2

$$\bar{t} = \sum_{i=1}^N \sum_{j=1}^N \frac{P_{ij}}{IP} \log_2 \left(\frac{D_{ij}}{W_j} + 1 \right) \quad (2)$$

เมื่อ \bar{t} คือ เวลาเฉลี่ยในการพิมพ์มีหน่วยเป็นวินาที
 P_{ij} คือ เมตริกซ์ค่าความน่าจะเป็นที่เกิดคู่อักษรใด ๆ
 D_{ij} คือ เมตริกซ์ระยะห่างระหว่างคู่อักษรใด ๆ
 W_j คือ ขนาดของแป้นอักษรปลายทาง
 IP คือ คำนวณประสิทธิภาพ
 N คือ จำนวนแป้นพิมพ์ทั้งหมด

อย่างก็ตามในการวัดประสิทธิภาพการพิมพ์บนแผงแป้นพิมพ์มีหน่วยวัดเป็น คำต่อนาที (word per minute, wpm) ดังนั้นสามารถคำนวณได้ดังสมการที่ 3

$$t_{wpm} = 60 / AWL \times \bar{t} \quad (3)$$

เมื่อ AWL คือ ความยาวเฉลี่ยต่อคำในคลังข้อความ
 \bar{t} คือ เวลาเฉลี่ยต่อคำ

3. วิธีการดำเนินงานวิจัย

งานวิจัยนี้ใช้ตัวอย่างคำจากโครงการวัดเปรียบเทียบสมรรถนะเพื่อพัฒนามาตรฐานการประมวลผลภาษาไทย (Benchmark for Enhancing the Standard for Thai

Language Processing, BEST) เลือกใช้คลังข้อความ (corpus) ชุดฝึกฝน (training set) ชุดที่ 1 จำนวน 5 ล้านคำ รวบรวมจาก 4 หมวดหมู่เอกสาร อันประกอบด้วย บทความ (article) ข่าว (news) นวนิยาย (novel) และสารานุกรม (encyclopedia) ซึ่งใช้ในการแข่งขัน BEST2010 จากนั้นนำมาวิเคราะห์หาความน่าจะเป็นของการเกิดของตัวอักษรแต่ละตัว และคู่อักษรด้วยเทคนิคการวิเคราะห์ความถี่เอ็น-แกรม (N-Gram) เพื่อนำผลลัพธ์ที่ได้มาใช้ในการคำนวณหาตำแหน่งที่เหมาะสมของแป้นพิมพ์ จากนั้นกำหนดขนาดของแป้นพิมพ์จำนวน 8 แถว 13 คอลัมน์ (8x13) กำหนดขนาดของแป้นพิมพ์ 100 จุดภาพ (pixel) และกำหนดพื้นที่ของแป้นพิมพ์ออกเป็น 7 ส่วนภายใต้แนวคิดการพิมพ์ นิ้วเดียวบนจอภาพสัมผัสผู้พิมพ์ต้องมองไปยังแป้นพิมพ์เสมือนซึ่งจุดสนใจของการมองของมนุษย์คือ จุดโฟกัสที่อยู่ตรงกลางของวัตถุ หรือกลางหน้าจอภาพระบบสัมผัสนั่นเอง

นอกจากแป้นพิมพ์อักษรแล้วงานวิจัยยังออกแบบตำแหน่งของแป้นควบคุมอีก 4 แป้น คือ Enter, Space, Back Space (BS) และ Switch Number (SN) เพื่อใช้เปลี่ยนโหมดระหว่างตัวเลขไทยและตัวเลขอารบิก โดยกำหนดตำแหน่งคงที่คือแป้นพิมพ์ควบคุมแป้นพิมพ์ Enter กำหนดตำแหน่งแถวที่ 7 คอลัมน์ที่ 13 แป้นพิมพ์ Space ตำแหน่งแถวที่ 7 คอลัมน์ที่ 12 แป้นพิมพ์ BS กำหนดตำแหน่งแถวที่ 8 คอลัมน์ที่ 12-13 และแป้นพิมพ์ SN กำหนดตำแหน่งแถวที่ 1 คอลัมน์ที่ 11 สำหรับอักษรตัวเลขไทย (๐-๙) และอารบิก (0-9) จำนวน 10 แป้น กำหนดตำแหน่งแน่นอนเช่นกันคือ แถวที่ 1 คอลัมน์ที่ 1-10

ขั้นตอนการคำนวณหาตำแหน่งแป้นพิมพ์ที่เหมาะสมด้วยการคำนวณเชิงวิวัฒนาการ เริ่มต้นโดยกำหนดให้โครโมโซม (chromosome) มีจำนวนยีนส์เท่ากับอักษร 89 ตัว ทั้งนี้ไม่นับรวมแป้นพิมพ์ตัวเลขและแป้นพิมพ์ควบคุมจำนวน 14 แป้นพิมพ์ ซึ่งอยู่ในกลุ่มพื้นที่ 7 โดยจัดเรียงเป็นกลุ่มจำนวน 6 กลุ่ม ซึ่งยีนส์แต่ละกลุ่มแทนข้อมูลอักษรตามพื้นที่ที่กำหนดไว้ มีขั้นตอนรายละเอียดดังนี้

- (1) สร้างประชากรโดยใช้วิธีการสุ่ม (random) จำนวน 100 ชุด เพื่อแทนโครงสร้างแผงพิมพ์
- (2) คำนวณเวลาในการพิมพ์แต่ละโครงสร้างแผงพิมพ์ด้วยสมการกฎของฟิต ทดสอบกับคำไทยที่ใช้บ่อยจำนวน

20,000 คำ จากนั้นนำผลลัพธ์ประเมินค่าความเหมาะสมด้วยวิธีสัดส่วน (proportional fitness function) เพื่อใช้ในการคัดเลือกโครโมโซมต้นกำเนิดสายพันธุ์ต่อไป

- (3) คัดเลือกสายพันธุ์ (selection) ต้นกำเนิดด้วยกลยุทธ์การคัดเลือกโครโมโซมที่โดดเด่นเพื่อรักษาไว้ในรุ่นถัดไป (elitist strategy) ร่วมกับวิธีวงล้อเสี่ยงทาย (roulette wheel selection)
- (4) ดำเนินการแลกเปลี่ยนยีนส์ (recombination) ด้วยการสลับสายพันธุ์ (crossover) โดยทั่วไปการสลับสายพันธุ์จะทำการสุ่มโครโมโซมจากประชากรมาสองโครโมโซม (พ่อและแม่) แล้วทำการแลกเปลี่ยนยีนส์ระหว่างโครโมโซม ซึ่งเป็นวิธีการเลียนแบบธรรมชาติสำหรับงานวิจัยนี้ ผู้วิจัยออกแบบการสลับสายพันธุ์ด้วยการนำแนวคิดการดัดแปลงทางพันธุวิศวกรรม (Genetically Modified Organisms, GMOs) ของพืชเพื่อให้ได้ผักหรือผลไม้ที่มีลักษณะพึงประสงค์ หลักการทำงานของกรสลับสายพันธุ์ด้วยวิธี GMOs คือ การแลกเปลี่ยนยีนส์จนได้โครโมโซมลูกหลานไม่ได้เกิดจากการสุ่มจากประชากรพ่อและแม่คู่เดียวเท่านั้น แต่เกิดจากการนำยีนส์เด่นที่สุ่มได้จากโครโมโซมประชากร ซึ่งลักษณะเด่นของยีนส์ขึ้นอยู่กับปัญหาของงานวิจัยแต่ละประเด็น และจำนวนครั้งที่สุ่มขึ้นอยู่กับจำนวนลักษณะเด่นที่กำหนด งานวิจัยนี้กำหนดให้ความน่าจะเป็นในการสลับสายพันธุ์ร้อยละ 80 ($P_c = 0.8$)
- (5) นำโครโมโซมลูกหลาน (offspring) ของเป็นพิมพ์แม่และแมงคำนวณหาเวลาในการพิมพ์ ด้วยสมการกฎของฟิต งานวิจัยนี้กำหนดให้ค่าดัชนีประสิทธิภาพ (IP) = 4.9 บิตต่อวินาที สำหรับกรณีพิมพ์ตำแหน่งซ้ำเมื่อ $i = j$ ตัวอย่างเช่น “วีรกรรม” กำหนดให้เวลาในการพิมพ์ตำแหน่งเป็นพิมพ์อักษร “ร” ซ้ำ ใช้เวลาในการพิมพ์ 0.127 วินาที
- (6) จากนั้นดำเนินการแทนที่โครโมโซมเป็นพิมพ์ลูกหลานแทนที่โครโมโซมเป็นพิมพ์เก่า (replacement) งานวิจัยนี้เลือกแทนที่ประชากรทั้งรุ่นเนื่องจากการคัดเลือกสายพันธุ์ด้วยกลยุทธ์การ

คัดเลือกโครโมโซมที่โดดเด่นเพื่อรักษาไว้ในรุ่นถัดไปไว้แล้ว

- (7) ดำเนินการทำซ้ำข้อ 2-6 จนเวลาที่ใช้ในการพิมพ์ของเป็นพิมพ์ที่คำนวณได้ไม่มีการเปลี่ยนแปลง หรือครบจำนวน 100 รุ่น ตามที่ได้กำหนดไว้ในงานวิจัยนี้

4. ผลวิจัย

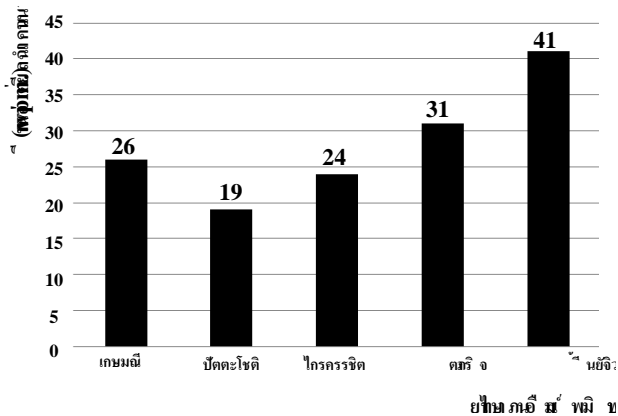
ผลการคำนวณหาตำแหน่งเป็นพิมพ์เสมือนภาษาไทยด้วยวิธีคำนวณเชิงวิวัฒนาการ โดยใช้โครงสร้างแมงพิมพ์จำนวน 100 แมง และดำเนินการคำนวณเชิงวิวัฒนาการจำนวน 100 รุ่น จากนั้นทำการทดสอบกับคำไทยที่ซับซ้อนได้จากการวิเคราะห์ความถี่จากคลังข้อความชุดฝึกฝน (training set) ชุดที่ 1 จำนวน 5 ล้านคำ รวบรวมจาก 4 หมวดหมู่เอกสาร ซึ่งใช้ในการแข่งขัน BEST2010 รวมทั้งสิ้นจำนวน 20,000 คำ ผลการทดลองพบว่า ได้เป็นพิมพ์เสมือนภาษาไทยใหม่ที่ทำให้ความเร็วเฉลี่ยในการพิมพ์เท่ากับ 41 คำต่อนาที และเวลาที่ใช้ในการพิมพ์ของเป็นพิมพ์ที่คำนวณได้ไม่มีการเปลี่ยนแปลงในรุ่นที่ 48 โครงสร้างเป็นพิมพ์เสมือนภาษาไทยใหม่แสดงดังภาพที่ 5

Ř	ř	Š	š	Ť	ť	Ť	ť	Ů	ů	SN	,	;
Ć	Ć	Ī	-	Ò	Ğ	Ý	À	Ĺ	ě	,	:	?
Ł	p	i	˘	²	Ž	ă	š	˘	ī	/	"	
ş	Ă	ą	ˆ	³	•	õ	İ	y	-	ğ	.	“
"	L	ž	˘	Ó	½	ć	Ş	˘	ı	-	(-
,	Š	Č	Đ	đ	ĩ	¾	Ě	Ī	×	Ú)	+
þ	ł	¼	Ů	ı		đ	Đ	Đ	ı	°	Space	Enter
ę	ł	Ř	õ	Ů	Ę	ř	F	Ů	č	ů	Backspace	

ภาพที่ 5: โครงสร้างแมงพิมพ์ผลลัพธ์ที่ได้จากงานวิจัยนี้

เพื่อเป็นการทดสอบประสิทธิภาพของโครงสร้างเป็นพิมพ์ใหม่ที่ได้จากงานวิจัยนี้ ผู้วิจัยดำเนินการเปรียบเทียบประสิทธิภาพกับโครงสร้างเป็นพิมพ์มาตรฐานเกษมณี เป็นพิมพ์แบบปัตตะโชติ โครงสร้างเป็นพิมพ์ของ จีรภาค และเกริก และโครงสร้างเป็นพิมพ์ของ ไกรครรชิต และเกริก ผลการเปรียบเทียบความเร็วเฉลี่ยในการพิมพ์ พบว่า โครงสร้างเป็นพิมพ์ของงานวิจัยนี้ให้ประสิทธิภาพสูงสุดที่ความเร็ว 41 คำต่อนาที และเป็นพิมพ์จากงานวิจัยของจีรภาค และเกริกให้ประสิทธิภาพรองลงมามีความเร็วในการพิมพ์ 31 คำต่อนาที

สำหรับเป็นพิมพ์มาตรฐานเกษมณี ให้ประสิทธิภาพที่ความเร็วในการพิมพ์ 26 คำต่อนาที ส่วนเป็นพิมพ์ปัตตะโชติ และเป็นพิมพ์ที่ได้จากงานวิจัยของไกรกรรชิต และเกริก ให้ประสิทธิภาพความเร็วในการพิมพ์ที่ 19 และ 24 คำต่อนาที ตามลำดับ รายละเอียดแสดงดังภาพที่ 6



ภาพที่ 6: ผลเปรียบเทียบประสิทธิภาพโครงสร้างเป็นพิมพ์เสมือน

5. บทสรุป

โครงสร้างเป็นพิมพ์ใหม่ที่คำนวณได้จากงานวิจัยนี้เป็นเป็นพิมพ์ชนิดหน้าเดียว ขนาด 8 แถว 13 คอลัมน์ และรูปทรงปุ่มพิมพ์เป็นสี่เหลี่ยม มีขนาดของแป้นพิมพ์ 100 จุดภาพ (pixel) ข้อดีของแป้นพิมพ์ชนิดหน้าเดียวคือไม่ต้องเสียเวลาในการสลับหน้า (switching pages) แต่สำหรับแป้นพิมพ์มาตรฐานแบบพิมพ์เกษมณี และแบบพิมพ์ปัตตะโชติ เป็นแป้นพิมพ์ชนิดสองหน้าทำให้เสียเวลาในการสลับหน้า โดยการคลิกที่ปุ่มชิฟ (Shift) เพื่อเปลี่ยนหน้าแป้นพิมพ์ ซึ่งถือเป็นข้อดีของแป้นพิมพ์ชนิดสองหน้า สำหรับแป้นพิมพ์ของจิรภาค และเกริก แม้เป็นแป้นพิมพ์ชนิดหน้าเดียวแต่ใช้ข้อความบนเครื่องข่ายทางสังคมออนไลน์เป็นข้อมูลหลักในการออกแบบซึ่งอาจทำให้ประสิทธิภาพของการพิมพ์ไม่สูงนักเมื่อใช้กับข้อมูลในกลุ่มอื่น และจากผลการเปรียบเทียบประสิทธิภาพโครงสร้างแป้นพิมพ์ แสดงให้เห็นว่าโครงสร้างแป้นพิมพ์ที่ได้จากผลการวิจัยในครั้งนี้มีประสิทธิภาพค่อนข้างสูงกว่าโครงสร้างแป้นพิมพ์ภาษาไทยดั้งเดิมที่มีอยู่ในปัจจุบัน อย่างไรก็ตามสิ่งหนึ่งที่ต้องตระหนักและคำนึงถึงคือ แป้นพิมพ์

เสมือนเป็นส่วนต่อประสานกับผู้ใช้งาน ดังนั้นการทดลองการใช้งานพิมพ์จริงกับผู้พิมพ์จึงเป็นงานที่ต้องวิจัยและพัฒนาต่อไป

เอกสารอ้างอิง

- [1] Zecevic, A., Miller, D., Harburn, K., "An evaluation of the ergonomics of three computer keyboards", *Ergonomics* 43(1), 55-72, 2000.
- [2] MacKenzie, I.S., Soukoreff, R.W., "Text entry for mobile computing: models and methods, theory and practice", *Human-Computer Interaction* 17, pp.147-198, 2002.
- [3] Zhai, S., B.A. Smith, and M. Munter, "Performance Optimization of Virtual Keyboard", *Human-Computer Interaction*, 17 (2&3), pp. 229-270. 2002.
- [4] Koanantakool, H. T., Karoonboonyanan, T., Wutiw WATCHAI, C., "Computer and the Thai Language", *Annals of the History of Computing, IEEE Journal*, vol.31 (1), 46-61, 2009.
- [5] S. Pattajoti, "The Evolution of The Typewriter", National Research Council, The Office of the Prime Minister, 4 Nov. 1966.
- [6] ไกรกรรชิต ดวงดี และเกริก ภิมย์โสภากา, "การวิเคราะห์และออกแบบแป้นพิมพ์เสมือนภาษาไทยสำหรับเครื่องคอมพิวเตอร์จอสัมผัสโดยใช้ขั้นตอนวิธีพันธุกรรม", งานประชุมวิชาการระดับชาติการไฟฟ้าอิล็กทรอนิกส์ คอมพิวเตอร์ โทรคมนาคม และสารสนเทศไทยแห่งประเทศไทย ครั้งที่ 12 (NCSEC2008), พัทยา, ประเทศไทย, 20-21 พฤศจิกายน 2551.
- [7] J. Simathamand and K. Piromsopa. (2011). "Performance Optimization of Thai Virtual Keyboard for Social Networking". In proceeding on Computer Science and Software Engineering (JCSSE 2011), 11-13 May. 2011, pp. 210-213.
- [8] Elaine Rich and Kevin Knight, *Artificial intelligence*, McGraw-Hill: New York, 1991.
- [9] Paul M. Fitts, "The information capacity of the human motor system in controlling the amplitude of movement", *Journal of Experimental Psychology*, volume 47, number 6 June 1954, pp.381-391.

เครื่องมือเพื่อบันทึกเวลาเหตุการณ์ที่เกิดจากการใช้งานของผู้ใช้บนวินโดวส์แอปพลิเคชัน

A User Action Events Timestamp Tool on Windows Application

รพีพงษ์ ศรีวิบูลย์ (Rapheephong Sriwiboon)¹ และ เชษฐ พัฒโนทัย (Chate Patanothai)²

ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

¹Rapheephong.s@student.chula.ac.th, ²chate.p@chula.ac.th

บทคัดย่อ

ในการทดสอบและประเมินผลความสามารถการใช้งานโปรแกรมประยุกต์ที่ถูกพัฒนาเสร็จแล้วนอกจากจะใช้วิธีการวัดผลด้วยการสังเกตโดยผู้สังเกตหรือการบันทึกวีดีโอแล้วในปัจจุบันมีเครื่องมือที่สามารถบันทึกการใช้งานของผู้ใช้แบบอัตโนมัติแต่ก็ยังต้องใช้งานภายใต้เงื่อนไขที่ผู้พัฒนาเครื่องมือได้กำหนดไว้เท่านั้นและไม่สามารถใช้กับโปรแกรมประยุกต์ทั่วไปได้ ในการใช้งานส่วนมากมักจะเป็นกลุ่มผู้พัฒนาซอฟต์แวร์เท่านั้นที่มักจะใช้เครื่องมือดังกล่าวเพื่อช่วยในการค้นหาข้อผิดพลาดที่เกิดขึ้นในซอฟต์แวร์

โครงการนี้จึงได้นำเสนอเครื่องมือเพื่อใช้ในการรวบรวมข้อมูลการใช้งานโปรแกรมสำหรับโปรแกรมประยุกต์ทั่วไปโดยใช้หลักการติดตามเหตุการณ์บนวินโดวส์ (Windows Events) และใช้กรอบการทำงาน UI Automation on .NET Framework เพื่อใช้ในการเข้าถึงคอมโพเนนต์ของโปรแกรมประยุกต์และเหตุการณ์ที่ต้องการบันทึก โดยประโยชน์ที่ได้จากเครื่องมือนี้คือทำให้ได้ข้อมูลการใช้งานแอปพลิเคชันจากผู้ใช้งานเพื่อใช้เป็นข้อมูลในการวิเคราะห์เพื่อเพิ่มความสามารถในการใช้งาน

คำสำคัญ: ยูสอะบิลิตี้ ยูไอโทมัสแตมปี ยูไอออโตเมชัน เหตุการณ์บนส่วนติดต่อผู้ใช้งาน การประเมินผล ส่วนติดต่อผู้ใช้งาน

Abstract

In usability testing and evaluation of developed application, apart from evaluation by observation and video recording, there are some tools that can be used to automatically record user event. But it must be used under constraint of developing condition and it cannot be

used with generic application. The developer only uses it for debugging his/her software.

In this project, we propose tool for collecting user event data on general application by using windows event monitoring and using UI Automation on .NET Framework for accessing UI component and events.

The benefit of this tool is to provide an actual data based usage of the developed application for improving usability.

Keywords: Usability, UI-Timestamp, UIAutomation, User Interface Event, User Interface Evaluation.

1. บทนำ

หนึ่งในปัจจัยคุณภาพที่มีความสำคัญและเป็นปัจจัยในการเลือกผลิตภัณฑ์ซอฟต์แวร์ของผู้ใช้งานคือความสามารถในการใช้งาน (Usability) ซึ่งเป็นแนวคิดที่เข้าใจและวัดได้ยากในการนิยามให้เฉพาะเจาะจง จึงต้องมีการสร้างและกำหนดการวัดและตัววัดที่เหมาะสมกับประสบการณ์และความสามารถของผู้ใช้งาน

เนื่องจากการวัดและการทดสอบความสามารถในการใช้งานผลิตภัณฑ์ซอฟต์แวร์ในรูปแบบดั้งเดิมที่ต้องการสังเกต จับเวลา และจดบันทึกนั้นต้องใช้ค่าใช้จ่ายและเวลาที่สูงรวมถึงความไม่แม่นยำของข้อมูลที่ได้จัดเก็บซึ่งอาจเกิดจากความผิดพลาดในการจดบันทึกข้อมูลหรือความประมาทของผู้ทดสอบ ที่มีผู้สังเกต หรือผู้จดบันทึกอยู่ด้วยขณะทำการทดสอบ

หนึ่งในขั้นตอนการออกแบบส่วนติดต่อผู้ใช้งานก็คือการรวบรวมและวิเคราะห์ข้อมูลการใช้งาน เครื่องมือและ

วิธีการต่างๆ ถูกนำมาใช้ในขั้นตอนการออกแบบ ส่วนติดต่อผู้ใช้งาน การเก็บข้อมูลการใช้งานมีได้หลากหลายวิธี เช่น การให้ผู้ใช้กรอกแบบสอบถาม การสัมภาษณ์ผู้ใช้ การให้ผู้ใช้ที่มีประสบการณ์ทดลองใช้เพื่อบันทึกข้อมูลเหตุการณ์ การบันทึกวิดีโอ และการบันทึกเหตุการณ์แบบอัตโนมัติ

วิธีการต่างๆ ที่กล่าวมานั้นมีการใช้งานที่แตกต่างกันและยังพบข้อจำกัดในหลายๆด้าน เช่น ผลลัพธ์ที่ได้ขึ้นอยู่กับตัวบุคคล ความถูกต้องและความแม่นยำในการบันทึกข้อมูล ใช้เวลาและทรัพยากรสูง [1] ได้ข้อมูลที่มีปริมาณมากเกินไปจนจำเป็นรวมถึงไม่ตรงตามความต้องการ และต้องใช้ผู้เชี่ยวชาญในการเก็บข้อมูล

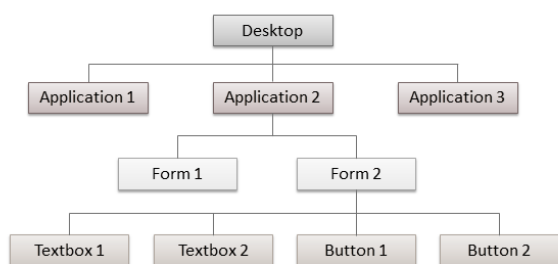
โครงการนี้จึงพัฒนาเครื่องมือเพื่อบันทึกเวลาเหตุการณ์ที่เกิดจากการใช้งานของผู้ใช้บนวินโดวส์แอปพลิเคชัน ให้ได้ข้อมูลที่เกิดจากผู้ใช้งานจริง มีความถูกต้องแม่นยำ ตรงตามความต้องการ รวมถึงใช้เวลาและทรัพยากรในปริมาณที่น้อยกว่าวิธีที่กล่าวมาข้างต้น

2. วรรณกรรมที่เกี่ยวข้อง

2.1 ยูสเซอร์อินเตอร์เฟซ ออโตเมชัน (UI Automation)

ยูสเซอร์อินเตอร์เฟซ ออโตเมชัน [2] เป็น คลาส (Class) ที่บรรจุอยู่ใน “System.Windows.Automation” ของ .NET Framework ที่ให้บริการการเข้าถึงข้อมูลส่วนประกอบแอปพลิเคชันที่อยู่บนส่วนติดต่อผู้ใช้งาน (User Interface, UI) บนหน้าจอเดสก์ท็อปด้วยการใช้ชุดคำสั่ง

ยูสเซอร์อินเตอร์เฟซ ออโตเมชันช่วยให้สามารถเข้าถึงข้อมูลของทุกคอมโพเนนต์ที่อยู่บนส่วนติดต่อผู้ใช้งานได้ โดยแอปพลิเคชันจะร้องขอผ่านทาง “AutomationElement” โดยที่องค์ประกอบทั้งหมดอยู่ภายในโครงสร้างต้นไม้ที่มีหน้าจอเดสก์ท็อปเป็นราก (root) ของโครงสร้างต้นไม้ ดังภาพที่ 1



ภาพที่ 1: โครงสร้างต้นไม้ของส่วนประกอบที่อยู่บนส่วนติดต่อผู้ใช้งาน

“AutomationElement” แสดงข้อมูลคุณสมบัติทั่วไปของคอมโพเนนต์ของแอปพลิเคชันที่อยู่บนส่วนติดต่อผู้ใช้งาน หนึ่งในคุณสมบัติเหล่านั้นก็คือประเภทของคอมโพเนนต์หรือคอนโทรล (control type) เช่นปุ่มกด (button) ซึ่งทำให้แอปพลิเคชันที่ร้องขอข้อมูลสามารถตัดสินใจได้ว่าจะมีปฏิสัมพันธ์อย่างไรกับคอมโพเนนต์นั้นๆ

ยูสเซอร์อินเตอร์เฟซ ออโตเมชันสามารถเฝ้าดูข้อมูลเหตุการณ์แบบเฉพาะบางเหตุการณ์และสามารถระบุเฉพาะในระดับคอมโพเนนต์ของแอปพลิเคชันเช่น กำหนดให้ติดตามเหตุการณ์เฉพาะเหตุการณ์ที่เกิดบนปุ่มบันทึกเท่านั้น

2.2 งานวิจัยที่เกี่ยวข้อง

มีงานวิจัยก่อนหน้าที่ได้ทำการเก็บรวบรวมข้อมูลเทคนิควิธีการในการจัดเก็บและการนำไปใช้สำหรับข้อมูลความสามารถในการใช้งานจากเหตุการณ์ที่เกิดขึ้นกับส่วนติดต่อผู้ใช้งานซึ่งสามารถแบ่งเป็นประเภทต่างๆได้ดังนี้ [3]

- การประสานเวลาและการค้นหา (synchronization and searching) คือการระบุตำแหน่งระหว่างเหตุการณ์ที่เกิดขึ้นกับสื่อประเภทอื่นเพื่อให้ง่ายในการค้นหาตำแหน่งของเหตุการณ์ที่เกิดขึ้นบนสื่อประเภทอื่น
- การเปลี่ยนรูป (Transformation) คือการเปลี่ยนข้อมูลเหตุการณ์ที่เกิดขึ้นให้เป็นข้อมูลที่ต้องการ
- การนับและสรุปข้อมูลสถิติ (Counts and Summary Statistics) คือการนับและการคำนวณข้อมูลทางสถิติเพื่อใช้อธิบายพฤติกรรมของผู้ใช้งาน
- การตรวจหาคำแหน่งการเกิดชุดลำดับของเหตุการณ์ (Sequence Detection) คือการตรวจหาคำแหน่งการเกิดชุดลำดับของเหตุการณ์เพื่อใช้ในการระบุปัญหาที่เกิดขึ้น
- การเปรียบเทียบชุดลำดับของเหตุการณ์ (Sequence Comparison) คือการเปรียบเทียบชุดลำดับที่ต้องการเพื่อใช้ในการหาข้อแตกต่างที่เกิดขึ้น เช่น การเปรียบเทียบชุดลำดับของเหตุการณ์ระหว่างผู้ใช้งานทั่วไปกับผู้ใช้งานที่มีประสบการณ์ เป็นต้น
- การแสดงผล (Visualization) คือการนำเอาข้อมูลเหตุการณ์มาแสดงผลในรูปแบบที่สามารถเข้าใจได้ง่ายเช่น การแสดงผลในรูปแบบกราฟ เป็นต้น

- การผสมผสาน (Integrated Support) คือการรวมรูปแบบต่างๆ เช่น การเปลี่ยนรูป การวิเคราะห์ และการแสดงผลเข้ามาไว้ด้วยกันเพื่ออำนวยความสะดวกในการใช้งานให้กับผู้ประเมิน

อีกหนึ่งงานวิจัยของ Janne Pitkänen, Matti Pitkäranta และ Marko Nieminen ได้นำเสนออุปกรณ์ดักจับข้อมูลการใช้งานที่ให้ชื่อว่า “UXBlackbox” [4] ซึ่งอุปกรณ์ดังกล่าวจะทำหน้าที่เป็นตัวกลางการสื่อสารข้อมูลระหว่างผู้ใช้งานกับแอปพลิเคชันที่ต้องการจะเก็บข้อมูลโดยการบันทึกข้อมูลนำเข้าที่เกิดขึ้นจากคีย์บอร์ดและเมาส์และยังมีอุปกรณ์นำเข้าพิเศษอีกหนึ่งอย่างคือปุ่ม “ชอบ” หรือ “ไม่ชอบ” ซึ่งข้อมูลนำเข้าทั้งหมดจะนำไปผนวกรวมกับการบันทึกวิดีโอในขณะที่ทำการทดสอบเพื่อให้ทราบว่าเหตุการณ์นำเข้าที่เกิดขึ้นนั้นเกิดขึ้น ณ เวลาใดบนวิดีโอที่ได้บันทึกไว้เพื่ออำนวยความสะดวกให้กับผู้ประเมิน

เนื่องจากเป็นอุปกรณ์พิเศษที่ใช้ในการบันทึกข้อมูลจึงทำให้ต้องทำการทดสอบภายในห้องทดลองที่ได้ติดตั้งอุปกรณ์เท่านั้นและไม่สามารถนำไปใช้ในวงกว้างได้เพราะต้องใช้อุปกรณ์จำนวนมาก

2.3 Keystroke-Level Model (KLM)

Keystroke-Level Model (KLM) [5] เป็นโมเดลที่ใช้ในการคาดการณ์เวลาที่จะต้องใช้เวลาในการดำเนินงานตามวิธีการที่กำหนดมีสูตรการคำนวณคือ $T_{\text{task}} = T_{\text{acquire}} + T_{\text{execute}}$ โดยที่ T_{execute} จะเป็นส่วนที่จะมีผลกับการออกแบบส่วนติดต่อผู้ใช้โดยตรงเนื่องจากเป็นเวลาที่ใช้ในการปฏิบัติงานกับระบบซึ่งมีสูตรการคำนวณคือ $T_{\text{execute}} = \sum (\text{time to execute primitive operation})$

2.4 เครื่องมือที่เกี่ยวข้อง

เครื่องมือที่สามารถติดตามเหตุการณ์ที่เกิดขึ้นกับส่วนติดต่อผู้ใช้งาน เช่น “Spy++” [6] เป็นยูทิลิตี้ที่สามารถติดตามการเกิดเหตุการณ์ที่เกิดขึ้นบนส่วนติดต่อผู้ใช้งานได้แต่ข้อมูลที่ให้เป็นข้อมูลที่แปลความได้ยากเนื่องจากเป็นข้อมูลระดับที่ระบบปฏิบัติการนำไปใช้และอีกหนึ่งเครื่องมือที่น่าสนใจก็คือ “UI Spy” [7] ผู้พัฒนาระบบและผู้ทดสอบมักจะนำไปใช้ในการตรวจสอบความถูกต้องของหน้าจอในขั้นตอนการทำงานแต่ไม่สะดวกในการใช้งานในด้านการติดตามเหตุการณ์ของผู้ใช้เพราะต้องใช้ผู้เชี่ยวชาญพอสมควรในการใช้งาน

จากงานวิจัยและเครื่องมือที่เกี่ยวข้องข้างต้นทำให้ผู้วิจัยได้แนวทางในการพัฒนาเครื่องมือเพื่อใช้ในการรวบรวมข้อมูลเหตุการณ์ที่เกิดขึ้นกับส่วนติดต่อผู้ใช้งานให้สามารถรวบรวมข้อมูลได้จากแอปพลิเคชันใดๆและเป็นข้อมูลที่ผู้ทดสอบหรือผู้ประเมินต้องการจริงๆเพื่อให้ผู้ประเมินหรือผู้ที่ต้องการข้อมูลดังกล่าวสามารถนำข้อมูลที่ได้รวบรวมได้นั้นไปใช้ประโยชน์ในการปรับปรุงส่วนติดต่อผู้ใช้งานให้เหมาะสมต่อไป

3. วิธีการดำเนินการวิจัย

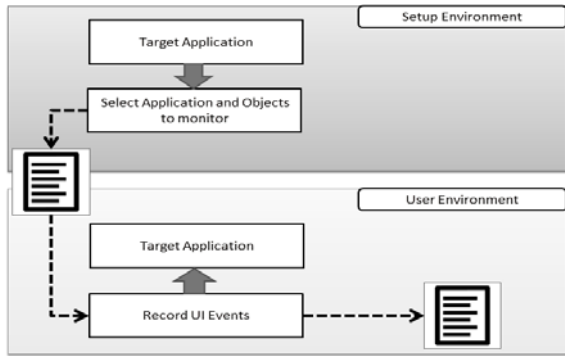
3.1 การวิเคราะห์ความต้องการของเครื่องมือเพื่อบันทึกเวลาเหตุการณ์ที่เกิดจากการใช้งาน

จากการวิเคราะห์ความต้องการเบื้องต้นของเครื่องมือบันทึกเวลาเหตุการณ์ที่เกิดจากการใช้งานของผู้ใช้บนวินโดวส์แอปพลิเคชันสามารถสรุปความต้องการได้ดังต่อไปนี้

- ระบบสามารถบันทึกข้อมูลเหตุการณ์ที่เกิดขึ้นบนส่วนติดต่อผู้ใช้งานได้โดยข้อมูลที่ถูกบันทึกจะประกอบด้วย คอมโพเนนต์ เหตุการณ์ และเวลาที่เกิด
- ระบบสามารถกำหนดหรือระบุเฉพาะคอมโพเนนต์ของส่วนติดต่อผู้ใช้งานที่ต้องการบันทึกเหตุการณ์
- ระบบสามารถบันทึกข้อมูลเหตุการณ์ของแอปพลิเคชันใดๆได้โดยไม่จำเป็นต้องมีการแก้ไขใดๆ
- ผู้ใช้ระบบไม่จำเป็นต้องมีความรู้หรือมีความเชี่ยวชาญในการโปรแกรมมิ่ง

3.2 การออกแบบและพัฒนาเครื่องมือบันทึกเวลาเหตุการณ์ที่เกิดจากการใช้งานของผู้ใช้บนวินโดวส์แอปพลิเคชัน

เครื่องมือบันทึกเวลาเหตุการณ์ที่เกิดจากการใช้งานของผู้ใช้บนวินโดวส์แอปพลิเคชันจะแบ่งการพัฒนาออกเป็น 2 ส่วน ดังแสดงได้ดังภาพที่ 2 คือ 1. ส่วนการเลือกส่วนประกอบและเหตุการณ์ของส่วนติดต่อผู้ใช้งานที่ต้องการบันทึกเหตุการณ์ (Setup Environment) 2. ส่วนการติดตามบันทึกผลส่วนประกอบและเหตุการณ์ของส่วนติดต่อผู้ใช้งานที่เลือก (User Environment)



ภาพที่ 2: โครงสร้างเครื่องมือบันทึกเวลาเหตุการณ์ที่เกิดจากการใช้งานของผู้ใช้บนวินโดวแอปพลิเคชัน

- ส่วนการเลือกส่วนประกอบและเหตุการณ์ของ ส่วนติดต่อผู้ใช้งานที่ต้องการบันทึกเหตุการณ์ (Setup Environment)

ในส่วนนี้ความสำคัญจะอยู่ที่การเลือกคอมโพเนนต์จากแอปพลิเคชันที่ทำงานอยู่ ซึ่งทางผู้พัฒนาได้ใช้ความสามารถในการดึงข้อมูลคอมโพเนนต์บนหน้าจอของ AutomationElement โดยการอ่านค่าตำแหน่งของเมาส์แล้วส่งไปให้เมธอด (Method) AutomationElement.FromPoint จะได้ออปเจ็กต์ AutomationElement ของคอมโพเนนต์ดังกล่าวมาใช้งานต่อไปได้ดังภาพที่ 3

```
System.Windows.Point point = new System.Windows.Point(Cursor.Position.X, Cursor.Position.Y);
AutomationElement element = AutomationElement.FromPoint(point);
```

ภาพที่ 3: ตัวอย่างการหาข้อมูลของคอมโพเนนต์บนหน้าจอ

- ส่วนการติดตามบันทึกผลส่วนประกอบและเหตุการณ์ของ ส่วนติดต่อผู้ใช้งานที่เลือก (User Environment)

ในส่วนนี้จะเป็นส่วนที่ใช้ในการตรวจสอบเหตุการณ์ที่เกิดขึ้นกับแอปพลิเคชันที่เราต้องการ โดยยูสเซอร์อินเตอร์เฟซอัตโนมัติเตรียมเมธอดไว้ให้ 4 เมธอดด้วยกันดังตารางที่ 1

ตารางที่ 1: เมธอดสำหรับการติดตามการเกิดเหตุการณ์

Method	Event Type
AddAutomationFocusChangedEventHandler	Focus change
AddAutomationPropertyChangedEventHandler	Property change
AddStructureChangedEventHandler	Structure change
	All other events, identified by an AutomationEvent
AddAutomationEventHandler	

จากภาพที่ 4 จะเป็นตัวอย่างที่ใช้ในการติดตามเหตุการณ์การถูกเรียกใช้งาน ซึ่งดูได้จาก การพารามิเตอร์

“InvokePattern.InvokeEvent” ของปุ่ม “elementButton” โดยที่เมื่อเหตุการณ์เกิดขึ้นทางยูสเซอร์อินเตอร์เฟซอัตโนมัติจะเรียกให้ฟังก์ชัน “OnUIAutomationEvent” ทำงานทำให้สามารถเก็บข้อมูลของเหตุการณ์ที่เกิดขึ้นได้

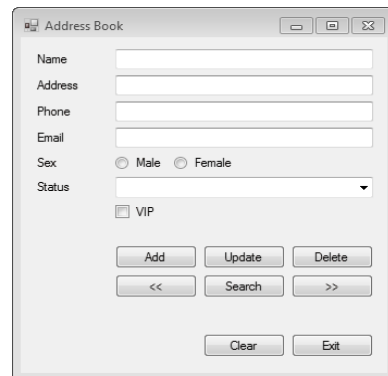
```
Automation.AddAutomationEventHandler(
    InvokePattern.InvokeEvent, elementButton, TreeScope.Element,
    UIEventHandler = new AutomationEventHandle(OnUIAutomationEvent));
```

ภาพที่ 4: ตัวอย่างการกำหนดให้ติดตามเหตุการณ์การเรียกใช้งานปุ่ม

4. ผลการดำเนินงาน

4.1 การทดสอบเครื่องมือและการประเมินผลการใช้งาน

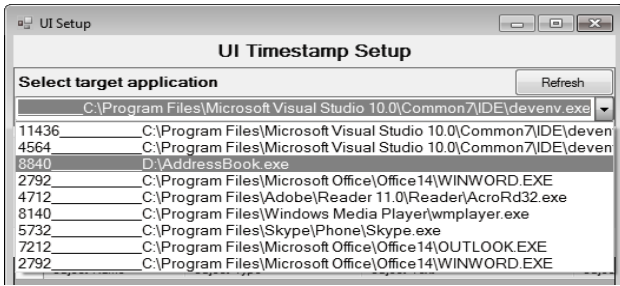
เนื่องจากต้องการทดสอบกับคอมพิวเตอร์ที่ที่หลากหลาย ผู้พัฒนาจึงได้ทำการออกแบบการทดสอบด้วยการสร้างแอปพลิเคชัน Address Book ดังภาพที่ 5 ขึ้นมาอย่างง่ายเพื่อใช้ในการทดสอบ



ภาพที่ 5: แอปพลิเคชัน Address Book

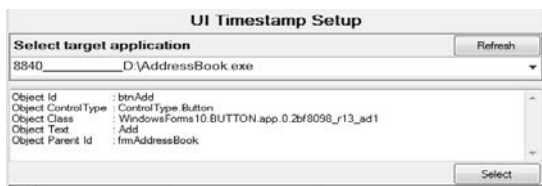
- การเปิดใช้งานเครื่องมือบันทึกเวลาเหตุการณ์
ความต้องการเบื้องต้นในการใช้งานเครื่องมือบันทึกเวลาเหตุการณ์คือต้องการ .NET Framework 4.0 (หรือสูงกว่า) และ windows XP เพื่อใช้ในการเปิดโปรแกรมขึ้นมาใช้งาน
- การใช้งานการกำหนดแอปพลิเคชันและการกำหนดคอมโพเนนต์

เมื่อเปิดแอปพลิเคชันในโหมดการกำหนดแอปพลิเคชันและคอมโพเนนต์ขึ้นมา โปรแกรมจะแสดงรายการของแอปพลิเคชันทั้งหมดขึ้นมาให้เลือกดังภาพที่ 6 เมื่อผู้ใช้เลือกแอปพลิเคชันที่ต้องการบันทึกข้อมูลแล้วผู้ใช้จะสามารถเลือกคอมโพเนนต์ภายใต้แอปพลิเคชันนั้นได้



ภาพที่ 6: โหมดการกำหนดแอปพลิเคชันและคอมพิวเตอร์

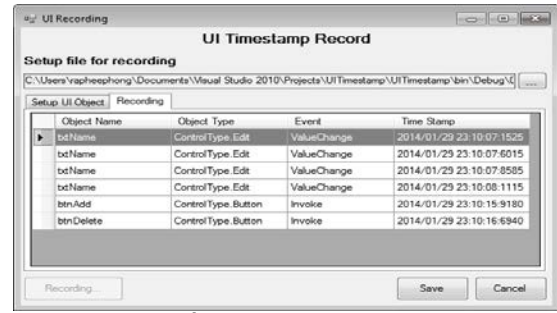
ซึ่งในการเลือกคอมพิวเตอร์ที่ต้องการบันทึกเหตุการณ์ที่เกิดขึ้นกับคอมพิวเตอร์นั้นจะเลือกจากแอปพลิเคชันจริงที่เปิดใช้งานอยู่ได้ในขณะนั้นด้วยการนำมาใส่ไปวางไว้บนคอมพิวเตอร์ที่ต้องการและกดปุ่ม Ctrl ตัวเครื่องมือจะแสดงกรอบสีแดงรอบคอมพิวเตอร์ดังกล่าว เพื่อแสดงให้ผู้ใช้ทราบว่านี่เป็นคอมพิวเตอร์ที่ต้องการหรือไม่ ในขณะเดียวกันที่หน้าจอเครื่องมือก็จะปรากฏข้อมูลของคอมพิวเตอร์ที่เลือกขึ้นมาให้เห็นดังภาพที่ 7



ภาพที่ 7: ข้อมูลของคอมพิวเตอร์ที่เลือก

- **การใช้งานการบันทึกข้อมูลเหตุการณ์ที่เกิดขึ้นบน แอปพลิเคชัน**

ในส่วนการใช้งานการบันทึกข้อมูลเหตุการณ์เป็นส่วนที่มีสำคัญมากเพราะต้องไม่ไปกระทบกับการใช้งานแอปพลิเคชันที่ต้องการทดสอบเนื่องจากอาจจะทำให้ได้ข้อมูลที่ไม่เป็นข้อมูลจากพฤติกรรมจริงของผู้ใช้งาน โดยการใช้งานการบันทึกข้อมูลเหตุการณ์จะเริ่มจากการเลือกไฟล์ที่ได้จากขั้นตอนการกำหนดแอปพลิเคชันและการกำหนดคอมพิวเตอร์และกดปุ่ม Record เท่านั้น หลังจากนั้นผู้ใช้งานก็สามารถใช้งานแอปพลิเคชันที่ต้องการทดสอบได้แบบปกติต่อไป โดยเมื่อมีเหตุการณ์ที่เกิดขึ้นกับคอมพิวเตอร์ที่ได้ถูกเลือกไว้ในขั้นตอนการกำหนดคอมพิวเตอร์ทางเครื่องมือก็จะแสดงข้อมูลเหตุการณ์ที่เกิดขึ้นในตารางดังภาพที่ 8 และเมื่อได้ทำการทดสอบหรือใช้งานแอปพลิเคชันเสร็จแล้วก็สามารถบันทึกข้อมูลเป็นไฟล์ .csv เพื่อนำไปใช้ในการวิเคราะห์ข้อมูลต่อไป



ภาพที่ 8: ข้อมูลเหตุการณ์ที่เกิดขึ้น

ผู้พัฒนาได้ทดลองกับแอปพลิเคชันตัวอย่างในงานเพิ่มผู้ติดต่อคนใหม่และงานค้นหาผู้ติดต่อเพื่อแก้ไขเบอร์โทรศัพท์เมื่อได้ผลลัพธ์แล้วจึงได้นำมาเปรียบเทียบกับการคำนวณโดยใช้ KLM ผลปรากฏว่าให้ผลลัพธ์ที่ใกล้เคียงกันคือ KLM สามารถได้ 10.9 และ 19 วินาที ตามลำดับ ในขณะที่เครื่องมือบันทึกเวลาสามารถบันทึกได้ 10.2 และ 17 วินาที ตามลำดับ ซึ่งจะเห็นว่าให้ผลลัพธ์ที่ใกล้เคียงกันแต่เมื่อผู้พัฒนาได้ทำการแก้ไขแอปพลิเคชันตัวอย่างให้ใช้งานได้ยากขึ้น เช่น ปรับขนาดปุ่มให้เล็กลงหรือวางในตำแหน่งที่ไม่เหมาะสมซึ่งผลจากการคำนวณค่า KLM จะได้เท่าเดิมเพราะยังคงมีลำดับการทำงานที่เหมือนเดิมแต่เมื่อนำมาทดลองกับเครื่องมือบันทึกเวลาสามารถบันทึกได้มากกว่า 10.2 และ 17 วินาที แสดงให้เห็นว่านอกจากเครื่องมือบันทึกเวลาจะสามารถวัดและประเมินเวลาในการทำงานได้เหมือนกับ KLM แล้วยังให้ข้อมูลที่เกิดจากการทำงานที่เป็นจริงมากกว่า

4.2 ปัญหาและข้อจำกัด

ถึงแม้ว่ายูสเซอร์อินเตอร์เฟซ ออโตเมชันจะสามารถเฝ้าดูเหตุการณ์ที่เกิดขึ้นบนส่วนติดต่อผู้ใช้งานได้แต่ยังมีข้อจำกัดในส่วนเหตุการณ์ที่สามารถเฝ้าดูได้ของแต่ละแอปพลิเคชันที่สร้างมาจากคนละกรอบการทำงานหรือแอปพลิเคชันเดียวกันแต่ทำงานบนระบบปฏิบัติการที่ต่างกันจะให้ผลของการเฝ้าดูเหตุการณ์ที่แตกต่างกันดังนั้นในขั้นต้นของโครงการนี้จึงรองรับแอปพลิเคชันที่สร้างจากกรอบการทำงาน .NET และทำงานบน windows XP เท่านั้น

5. สรุป

การทดสอบและประเมินความสามารถในการใช้งาน (Usability) เป็นส่วนสำคัญสำหรับแอปพลิเคชันในปัจจุบันเป็น

อย่างมากซึ่งปัญหาใหญ่ของการทดสอบและประเมินความสามารถในการใช้งานก็คือการได้มาของข้อมูลที่จะใช้ในการประเมินผล

เครื่องมือเพื่อบันทึกเวลาเหตุการณ์ที่เกิดจากการใช้งานของผู้ใช้บนวินโดวส์แอปพลิเคชันจึงเป็นเครื่องมือที่เป็นตัวช่วยผู้ประเมินในการรวบรวมข้อมูลการใช้งานจากผู้ใช้งานจริง แอปพลิเคชันจริง และใช้งานอยู่ในสภาพแวดล้อมจริง ซึ่งเป็นผลดีเป็นอย่างมากในการนำข้อมูลดังกล่าวไปวิเคราะห์เพื่อใช้ในการปรับปรุงส่วนติดต่อผู้ใช้ให้เหมาะสมต่อไป

เครื่องมือบันทึกเวลาที่นำเสนอเป็นเครื่องมือที่มีประสิทธิภาพทั้งด้านการให้ข้อมูลที่ต้องการ ข้อมูลที่ได้มีความแม่นยำและได้ข้อมูลในทันทีเมื่อใช้งานแอปพลิเคชันที่ต้องการทดสอบหรือวัดผลโดยสามารถเปรียบเทียบกับวิธีการอื่นๆดังตารางที่ 2

ตารางที่ 2: เปรียบเทียบการเก็บข้อมูลผู้ใช้งานแบบต่างๆ

วิธีการ	เวลาที่ใช้ในการเก็บข้อมูล (10 ผู้ใช้)	ปัญหาข้อจำกัด
การจับบันทึก	200 นาที	- ความแม่นยำในการบันทึกข้อมูล - จำนวนผู้จับบันทึกเท่ากับผู้ใช้
การอัดวีดีโอ	220 นาที	- เวลาที่ใช้อย่างต่ำเป็น 2 เท่าของงานที่ต้องการวัด (บันทึกและเล่นวีดีโอ) - จำนวนผู้แปลความจากวีดีโอเท่ากับผู้ใช้
การดักจับข้อมูลเมาส์และคีย์บอร์ด	> 20 นาที	- ข้อมูลที่ได้มีปริมาณมากและไม่ตรงตามความต้องการ จึงต้องนำไปแปลความหมายอีกขั้นก่อน
เครื่องมือบันทึกเวลาเหตุการณ์	20 นาที	

จากตารางที่ 2 ถ้าเวลาที่ใช้ในการทำงานให้เสร็จหนึ่งงานโดยประมาณเท่ากับ 20 นาทีและมีผู้รวบรวมข้อมูลผู้ใช้งานหนึ่งคนเท่านั้น จะเห็นว่าเครื่องมือบันทึกเวลาใช้เวลาโดยประมาณเท่ากับเวลาที่ใช้ในการทำงานปกติเท่านั้นในขณะที่การอัดวีดีโอจะใช้เวลามากที่สุดเพราะต้องใช้เวลา 20 นาทีสำหรับอัดวีดีโอและ 200 นาทีสำหรับการดูวีดีโอทั้งหมดเพื่อแปลงข้อมูลให้ออกมาในรูปแบบที่ต้องการ

ข้อมูลที่ได้จากการบันทึกเหตุการณ์ที่เกิดขึ้นสามารถนำมาใช้วิเคราะห์ข้อมูลทางตัวเลขได้หลากหลาย เช่น เวลาที่ใช้ในแต่ละงาน ความถี่ในการใช้งานคำสั่งใดๆ ความถี่ของเหตุการณ์ใดๆในช่วงเวลาใดๆ เป็นต้น

การวิเคราะห์ความถี่ของลำดับการทำงานของผู้ใช้ที่ได้จากการบันทึกเหตุการณ์อาจนำไปสู่การสร้างคำสั่งการทำงานที่

รวมเอาสองคำสั่งไว้ด้วยกันได้ หรืออย่างน้อยก็อาจเป็นการย้ายปุ่มคำสั่งมาไว้ใกล้ๆกันเพื่อเพิ่มความสะดวกในการทำงาน

ในลำดับต่อไปของโครงการนี้จะเป็นการนำเอาข้อมูลที่ได้มาทำการวิเคราะห์ให้เห็นว่าข้อมูลที่ได้มานั้นสามารถนำมาปรับปรุงความสามารถในการใช้งานได้ ในมุมมองใดบ้าง

ถึงแม้ว่าเครื่องมือเพื่อบันทึกเวลาเหตุการณ์ที่เกิดจากการใช้งานของผู้ใช้บนวินโดวส์แอปพลิเคชันจะสามารถรวบรวมข้อมูลจากผู้ใช้งานได้อย่างมีประสิทธิภาพก็ตามแต่ในส่วนของ การทดสอบและประเมินผลความสามารถในการใช้งานนั้น อาจจะยังไม่เพียงพอเนื่องจากความสามารถในการใช้งานยังมีส่วนสำคัญอีกหนึ่งประการนั้นคือความพึงพอใจของผู้ใช้งานซึ่งไม่สามารถวัดได้จากเครื่องมือเพื่อบันทึกเวลาจึงทำให้การทดสอบความสามารถในการใช้งานแบบดั้งเดิม เช่น การสังเกต ยังคงมีความสำคัญและควรได้รับการพัฒนาต่อไป

เอกสารอ้างอิง

- [1] M. Gray, A. Badre, and M. Guzdial. 1996. Visualizing usability log data. In Proceedings of the 1996 IEEE Symposium on Information Visualization (INFOVIS '96) (INFOVIS '96). IEEE Computer Society, Washington, DC, USA, 93-.
- [2] Microsoft MSDN Web Pages. URL: <http://msdn.microsoft.com/en-us/library/ms753107.aspx>, visited in 2013.
- [3] David M. Hilbert and David F. Redmiles. 2000. Extracting usability information from user interface events. ACM Comput. Surv. 32, 4 (December 2000), 384-421.
- [4] Janne Pitkänen, Matti Pitkäranta, and Marko Nieminen. 2012. Usability testing in real context of use: the user-triggered usability testing. In Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design (NordiCHI '12). ACM, New York, NY, USA, 797-798.
- [5] Card, S. K., Moran, T. P., and Newell, A. The keystroke-level model for user performance time with interactive systems. Commun. ACM 23, 7 (1980), 396-410.
- [6] Spy++ Web Pages. URL: [http://msdn.microsoft.com/en-us/library/aa242713\(v=vs.60\).aspx](http://msdn.microsoft.com/en-us/library/aa242713(v=vs.60).aspx), visited in 2013.
- [7] UI Spy Web Pages. URL: [http://msdn.microsoft.com/en-us/library/ms727247\(v=vs.90\).aspx](http://msdn.microsoft.com/en-us/library/ms727247(v=vs.90).aspx), visited in 2013.

การประเมินประสิทธิผลของส่วนต่อประสานแบบปรับเหมาะสมกับผู้ใช้ บนเว็ลด์ไวด์เว็บสำหรับผู้ใช้สายตาเลือนราง

An Assessment of Effectiveness of Adaptive Interfaces on WWW for Visually Impaired Users

พนิดา ฐัญญานนท์ (Panida Tunyanon)¹ และ วรสิทธิ์ ชูชัยวัฒนา (Worakit Choochaiwattana)²

สาขาวิชาวิศวกรรมเว็บ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยธุรกิจบัณฑิตย์

¹panida_tunyanon@hotmail.com, ²worakit.cha@dpu.ac.th

บทคัดย่อ

อินเทอร์เน็ตและเว็ลด์ไวด์เว็บเป็นช่องทางในการเผยแพร่ข้อมูลที่สำคัญในโลกสมัยใหม่ เว็บไซต์ส่วนใหญ่ถูกออกแบบมาสำหรับผู้ใช้งานทั่วไปเท่านั้น จากแนวคิดของเว็บไซต์ที่ทุกคนเข้าถึงได้ การออกแบบส่วนต่อประสานที่มีความสามารถให้การปรับเปลี่ยนการแสดงผล ให้เหมาะสมกับลักษณะของผู้ใช้งาน จึงเป็นเรื่องที่ท้าทาย งานวิจัยชิ้นนี้ เสนอแนวทางการสร้างส่วนต่อประสานแบบปรับเหมาะสมกับผู้ใช้บนเว็ลด์ไวด์เว็บสำหรับผู้ใช้ที่มีสายตาเลือนราง โดยพิจารณาจากลักษณะการมองเห็นของผู้ใช้ที่มีสายตาเลือนราง การประเมินประสิทธิผลของแนวทางนั้น สมาชิกของสมาคมคนสายตาเลือนรางแห่งประเทศไทยจำนวน 40 คนได้รับเชิญเข้ามาทดสอบ โดยใช้วิธีสุ่มตัวอย่างแบบเจาะจง การประเมินถูกแบ่งออกเป็นสองส่วน ได้แก่ การใช้แบบทดสอบความสามารถในการเข้าถึงเนื้อหาโดยการกำหนดงานให้ผู้ทดลองได้ปฏิบัติ และการใช้แบบสอบถามความคิดเห็นการใช้ประโยชน์เว็บ จากผลการทดลองพบว่าระยะเวลาในการทำงานที่ได้รับมอบหมายของผู้ใช้ และจำนวนครั้งของการคลิกในระบบนำทาง เมื่อใช้เว็บที่ส่วนต่อประสานแบบปรับเหมาะสมกับใช้น้อยกว่าที่ไม่ส่วนต่อประสานแบบปรับเหมาะสมกับผู้ใช้ อย่างมีนัยสำคัญทางสถิติ และจากแบบสอบถามความคิดเห็นการใช้ประโยชน์เว็บ พบว่าผู้ใช้มีความพึงพอใจอยู่ในระดับดี

คำสำคัญ: ส่วนต่อประสานแบบปรับเหมาะสมกับผู้ใช้
เว็บไซต์ที่ทุกคนเข้าถึงได้ ผู้มีสายตาเลือนราง

Abstract

Internet and World Wide Web has become the most important information dissemination channel in the modern world. However, most websites currently on the Internet are designed for general internet users. From the Web Accessibility concepts, tasks of designing adaptive user interfaces are challenging. This paper proposed guideline for developing adaptive interfaces on WWW for visually impaired user. These proposed guidelines consider their visual characteristics. In order to assess an effectiveness of the proposed guidelines, 40 visually impaired people from Low Vision Association of Thailand were invited to subjects in the experiment using purposive sampling technique. The assessment is divided into two parts, which are a web accessibility testing by tasks given and a web usability evaluation questionnaire. The results of the experiment showed that, when the subjects using adaptive interface, an average of time required for given tasks and an average number of clicks in website navigation are statistical significant lower than when the subjects using non adaptive interface. In addition, the results from the questionnaire showed that the subjects were satisfied with the adaptive interfaces.

Keyword: Adaptive User Interface, Web Accessibility, Visually Impaired People.

1. บทนำ

ในปัจจุบันเว็บไซต์ได้กลายเป็นสื่อที่มีความสำคัญและมีบทบาทอย่างมากในการให้ข้อมูลข่าวสารและประชาสัมพันธ์ ทั้งนี้เนื่องจากอินเทอร์เน็ตเป็นระบบเครือข่ายเปิดและสามารถติดต่อเชื่อมโยงตลอด 24 ชั่วโมง ดังนั้นบริษัท ห้างร้าน สถานประกอบการเกือบทุกแห่งทั้งภาครัฐและภาคเอกชนต่างจัดทำเว็บไซต์ขึ้นเพื่อเป็นช่องทางการสื่อสารในการเผยแพร่ข้อมูลข่าวสารไปสู่กลุ่มเป้าหมายโดยการติดต่อสื่อสารระหว่างกันเกิดความรวดเร็วและไม่ใช้เรื่องยากสำหรับบุคคลทั่วไปไปโดยยังมีบุคคลอีกกลุ่มที่มีลักษณะทางกายภาพอาจเป็นอุปสรรค สำหรับประเทศไทยนั้นได้มีกรอบนโยบายเทคโนโลยีสารสนเทศ ระยะ พ.ศ. 2554-2563 ซึ่งมีกำหนดวิสัยทัศน์ในการพัฒนาสังคมใจความว่า คนไทยทุกคนสามารถเข้าถึงและใช้ประโยชน์จากเทคโนโลยีสารสนเทศได้อย่างทั่วถึงและเท่าเทียมกัน จึงมีการจัดทำรูปแบบการพัฒนาเว็บไซต์ให้เป็นเว็บไซต์ที่ทุกคนสามารถเข้าถึง และกำหนดเกณฑ์มาตรฐานฉบับภาษาไทยขึ้นภายใต้ชื่อ “Thai Web Content Accessibility Guidelines” (TWCAG) เป็นการนำ WCAG 2.0 มาปรับใช้อันเป็นแนวทางในการส่งเสริมให้หน่วยงานต่างๆปรับปรุงและพัฒนาเว็บไซต์ที่ผู้ด้อยโอกาส ผู้สูงอายุและคนพิการ สามารถเข้าถึงได้ในการรับข้อมูลสารสนเทศ จากการสำรวจสถานะตาบอดแห่งชาติพบว่าอัตราผู้พิการสายเคเบิลในประชากรร้อยละ 0.81 และมีแนวโน้มที่จะเพิ่มขึ้น เนื่องจากผู้ที่มีสายเคเบิลมองไม่ได้ตาบอดสนิท สามารถดำเนินชีวิตอยู่ร่วมในสังคมและสื่อสารกับทั่วไปได้อย่างปกติและสามารถใช้งานคอมพิวเตอร์ในการหาข้อมูลข่าวสารบนเว็บไซต์แต่บางคนยังจำเป็นที่จะต้องใช้อุปกรณ์ช่วยเหลือพิเศษ เช่น โปรแกรมขยายหน้าจอ แต่ก็ยังไม่สามารถเข้าถึงข้อมูลข่าวสารบนเว็บไซต์ได้อย่างสะดวก เนื่องจากผู้พิการสายเคเบิลมีการมองเห็นที่แตกต่างกันไปตามลักษณะอาการและประเภทของโรคที่ทำให้เกิดสายเคเบิลมอง จากแนวคิดเห็นปัญหาดังกล่าวบทความนี้จึงได้นำเสนอการพัฒนาการพัฒนาส่วนต่อประสานแบบปรับเหมาะกับผู้ใช้ในการเข้าถึงสำหรับผู้ใช้งานเว็บผู้พิการทางสายตา ประเภทผู้มีสายเคเบิลมอง ซึ่งรวมถึงการออกแบบส่วนต่อประสานกับผู้ใช้งานที่เป็นส่วนที่ติดต่อระหว่างผู้ใช้งานเว็บไซต์ เพื่อรองรับการนำข้อมูลหรือคำสั่งเข้าสู่เมนู ตลอดจนนำเสนอสารสนเทศกลับมา

ยังผู้ใช้งานให้ปรับปรับตามความแตกต่างของผู้ใช้เว็บผู้พิการทางสายตา ประเภทผู้มีสายเคเบิลมองโดยอัตโนมัติ

2. วรรณกรรมและงานวิจัยที่เกี่ยวข้อง

2.1 ส่วนต่อประสานแบบปรับเหมาะกับผู้ใช้งาน

ส่วนต่อประสานแบบปรับเหมาะกับผู้ใช้งาน (Adaptive User Interface) ได้รับรูปแบบของการศึกษาและการวิเคราะห์เกี่ยวกับการปฏิสัมพันธ์ระหว่างมนุษย์และอุปกรณ์คอมพิวเตอร์โดยมีจุดประสงค์หลักเพื่อการใช้งานที่ง่ายและสะดวกต่อผู้ใช้งานหรือที่เรียกว่า Human-Computer interaction มาเป็นเวลานาน ซึ่งมีคุณลักษณะสำคัญของความฉลาดของส่วนต่อประสานแบบปรับเหมาะกับผู้ใช้งาน (Intelligent User Interface) [1, 2] ส่วนต่อประสานแบบปรับเหมาะกับผู้ใช้งานเป็นไปตามหลักการที่ได้กำหนดไว้โดยการประดิษฐ์หรือคิดค้นซอฟต์แวร์ ที่มีความสามารถปรับเปลี่ยนได้โดยอัตโนมัติมีการปรับเปลี่ยนตามลักษณะความสามารถของคอมพิวเตอร์ (Functionality) หรือส่วนต่อประสาน และยังสามารถการปรับปรุงและปรับเปลี่ยนตามผู้ใช้งาน [3, 4]

งานวิจัยของ Talia Lavie [5] ได้ทำการศึกษาประโยชน์และค่าใช้จ่ายของส่วนต่อประสานแบบปรับเหมาะกับผู้ใช้งาน โดยนำเสนอผลกระทบทั้งแง่บวกและแง่ลบของสื่อประสานแบบปรับเหมาะกับผู้ใช้งานในระบบ Telematic ซึ่งเป็นฟังก์ชันการทำงานของรถยนต์ ซึ่งมีการทดสอบการใช้ส่วนต่อประสานแบบปรับเหมาะสมมีการทดสอบโดยการกำหนดงานแก่ผู้ทดลองได้ปฏิบัติที่แตกต่างกันไปเช่น อายุ ระดับความยากของผลการทดสอบพบว่าผู้ใช้งานส่วนต่อประสานแบบปรับเหมาะสามารถทำงานได้ตามเป้าหมายของงานที่ผู้วิจัยได้กำหนดขึ้น

2.2 เว็บไซต์ที่ทุกคนเข้าถึงได้

เว็บไซต์ที่ทุกคนเข้าถึงได้ [6] ในภาษาอังกฤษที่เรียกว่า Web Accessibility หมายถึง กลุ่มคนทุกกลุ่มที่มีความหลากหลาย ไม่ว่าจะเป็นสภาพความเป็นอยู่ สภาพร่างกาย จิตใจ ผู้พิการหรือทุพพลภาพ ผู้สูงอายุ จะต้องครอบคลุมกลุ่มผู้ใช้งานทุกกลุ่ม ให้สามารถเข้าถึงข้อมูล สามารถรับรู้ เข้าใจ สิ่งงานหรือมีความต้องการใช้งานได้อย่างเท่าเทียมกัน โดยการเข้าถึงข้อมูลเหล่านั้น อาจจะใช้วิธีที่แตกต่างกันไป ตามวิธีการใช้งานของแต่ละกลุ่มผู้ใช้ เช่น ผู้พิการทางสายตา จะใช้โปรแกรมอ่าน

หน้าจอ หรือ โปรแกรมอ่านออกเสียงบนหน้าจอ โปรแกรมขยายหน้าจอ เพื่อให้สามารถอ่านได้ชัดเจนง่ายขึ้น หรืออุปกรณ์พกพา สามารถเข้าถึงข้อมูลได้เหมือนกับผู้ใช้คอมพิวเตอร์ทั่วไป

งานวิจัยของ จุฬารัตน์ แซ่เนี้ยว [7] นำเสนอการพัฒนาเครื่องมือสำหรับการแปลงหน้าเว็บเพจให้เป็นไฟล์สำหรับเว็บที่ทุกคนเข้าถึงได้โดยยึดแนวทางในการพัฒนาเว็บไซต์ให้ทุกคนเข้าถึงตามมาตรฐาน WCAG 1.0 โดยมุ่งเน้นไปยังกลุ่มคนพิการทางการมองเห็น ผลการทดสอบระบบทั้งในด้านความสามารถในการเข้าใช้งานและฟังก์ชันการทำงานของระบบอยู่ในระดับดี

2.3 ผู้มีสายตาเลือนราง

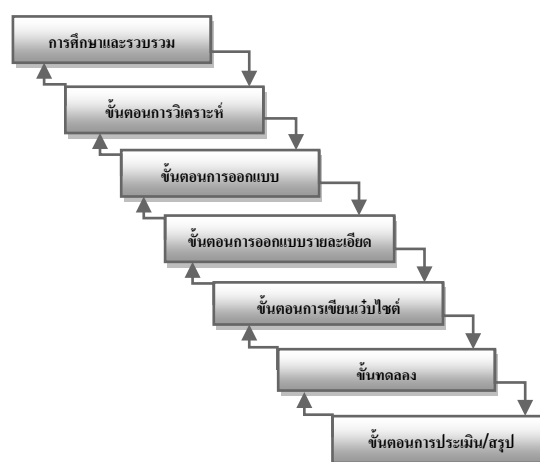
ตามกฎกระทรวงฉบับที่ 2 (พ.ศ.2537) ออกตามความในพระราชบัญญัติการฟื้นฟูสมรรถภาพคนพิการปี พ.ศ. 2534 คนพิการทางการมองเห็น ได้แก่ คนที่มีสายตาข้างที่ดีกว่าเมื่อใช้แว่นสายตาธรรมชาติแล้วมองเห็นน้อยกว่า 6/18 หรือ 20/70 ลงไปจนมองไม่เห็นแม้แต่แสงสว่างหรือคนที่มีลานสายตาแคบกว่า 30 องศาตามเกณฑ์ ICD-9-CM สายตาเลือนราง หมายถึง สายตา 20/80/-20/1000 และตาบอด หมายถึงสายตา 20/150 จนถึงมองไม่เห็นแสง ถ้าใช้เกณฑ์องค์การอนามัยโลก คือ สายตาน้อยกว่า 20/400 แต่ในสหรัฐอเมริกาใช้เกณฑ์น้อยกว่าหรือเท่ากับ 20/200 โดยสามารถแบ่งประเภทผู้มีสายตาเลือนรางตามประเภทของโรคที่ทำให้เกิดสายตาเลือนรางดังนี้ [8]

1. โรคต้อกระจก เป็นภาวะที่ใช้เรียกเลนส์ตาที่มีความขุ่นเกิดขึ้น อาจมีสีขาวขุ่น สีเหลือง ทำให้แสงผ่านเข้าไปยังจอประสาทตาได้น้อยลงภาพที่เกิดขึ้นจึงไม่ชัดเจน เกิดอาการที่เรียกว่าตามัว
2. โรคสายตาสั้น เป็นภาวะที่แสงผ่านกระจกตาและเลนส์ตา มาโฟกัสหน้าจอประสาทตา ทำให้ภาพที่ตกบนจอประสาทตาไม่ชัดเจน ทำให้แสงที่ผ่านกระจกตาและเลนส์ตา มาโฟกัสหน้าจอประสาทตา ภาพที่เห็นจึงไม่คมชัด
3. โรคจุดรับภาพเสื่อมหรือโรคศูนย์กลางจอประสาทเสื่อมจากอายุ เป็นสาเหตุสำคัญที่ทำให้ผู้สูงอายุ สูญเสียการมองเห็นอย่างถาวร อาการ ตาพร่ามัว มองเห็นตรงกลางของภาพไม่ชัดเจน หรือมองเห็นภาพบิดเบี้ยว

4. โรคต้อหิน ลักษณะอาการของผู้ป่วยโรคต้อหิน จะเริ่มจากสูญเสียการมองเห็นของลานสายตาครอบนอกก่อน เมื่อการดำเนินของโรครุนแรงมากขึ้นลานสายตาจะแคบลงเรื่อยๆ และตาบอดในที่สุดหากไม่ได้รับการรักษาที่เหมาะสมทันที

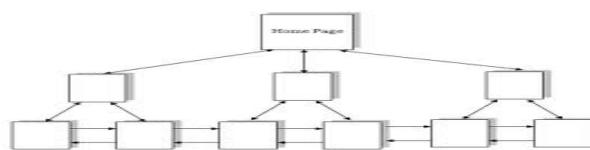
3. วิธีการดำเนินงานวิจัย

งานวิจัยนี้เสนอแนวทางของการสร้างส่วนต่อประสานแบบปรับเหมาะสมกับผู้ใช้บนเว็ลด์ไวด์เว็บสำหรับผู้ใช้ที่มีสายตาเลือนราง โดยพิจารณาจากลักษณะการมองเห็นของผู้ใช้ที่มีสายตาเลือนราง วิธีการวิจัยได้ประยุกต์การดำเนินงานตามทฤษฎี Waterfall Model ดังภาพที่ 1



ภาพที่ 1: แบบจำลองน้ำตก (Waterfall Model)

แบ่งออกเป็น 7) **ศึกษาและเก็บรวบรวมข้อมูล** เกี่ยวกับการออกและพัฒนาส่วนต่อประสานแบบปรับเหมาะสมกับผู้ใช้ในการเข้าถึงเว็บไซต์ สำหรับผู้มีสายตาเลือนราง 2) **ขั้นวิเคราะห์การออกแบบตามเอกลักษณ์และการรับรู้เข้าใจ** ของแต่ละประเภทของโรคสายตาเลือนราง 3) **ขั้นตอนการออกแบบ และพัฒนา** ในขั้นตอนนี้ผู้วิจัยออกแบบโดยอาศัยแนวทาง TWCA2010 [10] โดยทำการออกแบบเว็บไซต์ที่ไม่ได้ทำการปรับเหมาะสมกับผู้ใช้ (Non Adaptive Version) และได้ทำการปรับเหมาะสมกับผู้ใช้ (Adaptive Version) โดยมีเนื้อหาและโครงสร้างเว็บไซต์รวมทั้งระบบนำทางที่เหมือนกัน โดยใช้การออกแบบโครงสร้างแบบลำดับขั้น ดังภาพที่ 2



ภาพที่ 2: โครงสร้างเว็บไซต์แบบลำดับขั้น

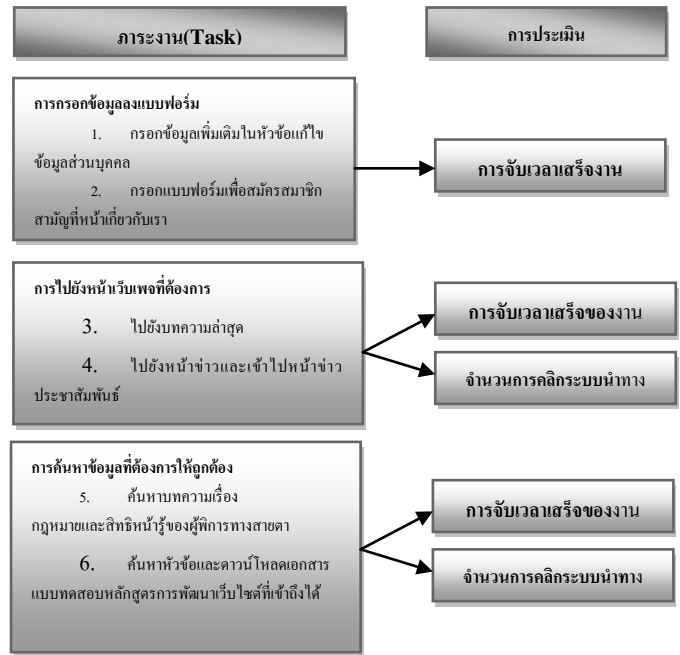
จะแตกต่างกันแค่ส่วนการออกแบบส่วนต่อประสานแบบปรับ
เหมาะสมกับผู้ใช้ โดยพิจารณาจากลักษณะการมองเห็นของ
ผู้ใช้ที่มีสายตาเลือนราง แบ่งออกได้เป็น 4 กลุ่ม ได้แนวคิดการ
ออกแบบจากงานวิจัย [9, 10] ตามตารางที่ 1

ตารางที่ 1: การออกแบบส่วนต่อประสานแบบปรับเหมาะสมโดยแบ่ง
ตามผู้ใช้งานตามประเภทผู้มีสายตาเลือนราง

ประเภทผู้มีสายตา เลือนราง	การปรับเปลี่ยนตามผู้ใช้
โรคต้อกระจก(Cataract)	<ul style="list-style-type: none"> มีการปรับขนาดตัวอักษรและนำหน้าของตัวอักษร 4 เท่าของขนาดปกติ มีการเปลี่ยน Contrast หรือการเปลี่ยนสีให้เป็นสีโทนมืด ขนาดหน้า 1024*768
โรคสายตาสั้น(Myopia)	<ul style="list-style-type: none"> มีการปรับขนาดตัวอักษรให้มีขนาดใหญ่ขึ้นกว่าเดิม ความหนาตัวอักษรเท่าเดิม ขนาดหน้าจอใช้ 1024*768
โรคจุดรับภาพเสื่อม (Macular Degeneration)	<ul style="list-style-type: none"> มีการปรับขนาดตัวอักษรให้มีขนาดใหญ่ขึ้นกว่าเดิม มีการปรับ Type Face หรือแบบอักษรจากปกติโดยใช้ชุดตัวอักษร DB Fongnamas มีการปรับขนาดรูปให้ใหญ่ขึ้น มีการเปลี่ยน Contrast หรือการเปลี่ยนสีให้เป็นสีโทนมืด ขนาดหน้าจอใช้ 1024*768
โรคจุดรับภาพเสื่อม (Macular Degeneration)	<ul style="list-style-type: none"> มีการปรับขนาดตัวอักษรให้มีขนาดใหญ่ขึ้นกว่าเดิมและหนาขึ้นกว่าเดิม ปรับขนาดหน้า(Resolution)เป็น 800*600 ปรับปุ่ม Cursor ให้ใหญ่ขึ้นกว่าเดิม

4) **ขั้นตอนการทดลอง** ในการทดลองนั้นผู้วิจัยจะทำการทดลองโดยใช้เว็บไซต์ที่ไม่ได้การปรับเหมาะสมกับผู้ใช้และทดลองอีกครั้งโดยใช้เว็บไซต์ที่ได้ทำปรับเหมาะสมกับผู้ใช้ที่มีการปรับเปลี่ยนตามกลุ่มผู้มีสายตาเลือนรางออกเป็น 4 กลุ่มตามประเภทของโรคสายตาเลือนราง 5) **ขั้นตอนการประเมินของผู้ใช้งานและสรุป** ในขั้นตอนนี้จะทำการประเมินนี้ผู้วิจัยจะแบ่งการประเมินออกเป็น 2 ส่วนคือ

1. การประเมินการเข้าถึงเนื้อหาโดยการกำหนดงาน (Task) ประเมินจากขั้นตอนการทดลองที่ผู้วิจัยได้กล่าวไปแล้ว โดยให้ผู้ทดลองปฏิบัติงานตามที่ได้กำหนดทั้งหมด 6 งาน แยกออกเป็น 3 ประเด็นเพื่อให้ครอบคลุมการทดสอบการใช้งานเว็บไซต์ โดยในแต่ละประเด็นจะมีความแตกต่างในการประเมิน ดังภาพที่ 3



ภาพที่ 3: ภาพแสดงการใช้หลักการประเมินผลโดยการกำหนดงาน

สำหรับการทดสอบโดยการกำหนดงาน (Task) นั้นผู้วิจัยได้พิจารณาจากลักษณะของการใช้งานของผู้ใช้เว็บซึ่งเป็นผู้มีสายตาเลือนราง และนำมาออกแบบโดยเป็นการกำหนดงาน สำหรับการทดลองโดยแบ่งงานออกเป็นทั้งหมด 6 งานให้ผู้ทดลองใช้งานเว็บไซต์ที่ไม่ได้ทำการปรับเหมาะสมกับผู้ใช้และเว็บไซต์ที่ได้ทำการปรับเหมาะสมกับผู้ใช้ เพื่อทำการเปรียบเทียบและประเมินผลและเพื่อลดผลกระทบที่เกิดจากการเรียนรู้ (Learning Effects) ผู้ทดลองแต่ละท่านจะเริ่มใช้งานเว็บไซต์สลับกัน

2. การประเมินความสามารถในการใช้งาน (Usability) ได้ทำการประเมินโดยสมาชิกสมาคมสายตาเลือนรางแห่งประเทศไทยจำนวน 40 คนโดยวิธีสุ่มตัวอย่างแบบเจาะจง เครื่องมือที่ใช้ในงานวิจัยคือ แบบสอบถามความคิดเห็นการใช้ประโยชน์จากเว็บโดยอาศัยแนวคิดจาก ISO 9241-11 เป็นการทดสอบเพื่อการยอมรับโดยผู้ใช้งาน User Acceptance Testing (UAT) โดยมีลักษณะคำถามให้ครอบคลุมคุณลักษณะ 4 ด้านคือ ประสิทธิภาพการใช้งาน (Effectiveness) ประสิทธิภาพในการใช้งาน (Efficiency) ความง่ายต่อการเรียนรู้ (Learnability) และความพึงพอใจในการใช้งาน (Likeability) เพื่อประเมินความคิดเห็นการใช้ประโยชน์จากเว็บโดยผู้วิจัยได้สร้างแบบประเมินมาตราส่วน (Rating Scale) ชนิด 5 ระดับตามวิธีของลิเคิร์ต (Likert) ในการตรวจให้คะแนนการวัดได้กำหนดเกณฑ์การให้คะแนนตามเกณฑ์ต่อไปนี้

	ข้อความที่เป็นบวก	ข้อความที่เป็นลบ
เห็นด้วยอย่างยิ่ง	5	1
เห็นด้วย	4	2
ไม่แน่ใจ	3	3
ไม่เห็นด้วย	2	4
ไม่เห็นด้วยอย่างยิ่ง	1	5

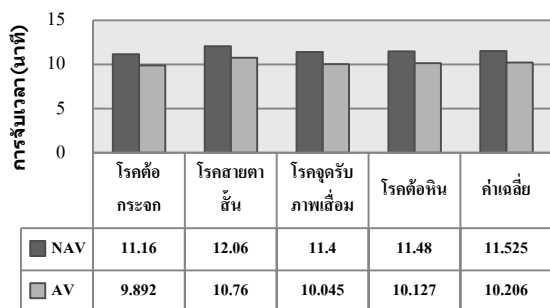
4. ผลการดำเนินงาน

4.1 ผลการทดสอบความสามารถระบบว่ามีความง่ายต่อการใช้งานและการเข้าถึงเนื้อหาโดยการกำหนดงาน(Task) ให้ผู้ทดลองได้ปฏิบัติคำสั่ง ของเว็บไซต์ที่ไม่ได้ทำการปรับเหมาะกับผู้ใช้งาน (NAV) และเว็บไซต์ที่ได้ทำการปรับเหมาะกับผู้ใช้งาน (AV) ดังนี้

ตารางที่ 2: แสดงการเปรียบเทียบผลการทดสอบโดยใช้การประเมินการจับเวลาเสร็จงาน (นาที) ของเว็บไซต์ที่ไม่ได้ทำการปรับเหมาะกับผู้ใช้งาน และเว็บไซต์ที่ได้ทำการปรับเหมาะกับผู้ใช้งาน

เวอร์ชัน	จำนวน N	ค่าเฉลี่ย \bar{x}	ค่า S.D	ค่า t	Sig.
NAV	40	11.525	1.91625	13.027	.000
AV	40	10.206	1.91693		

จากตารางที่ 2 พบว่า ผลการทดสอบจับเวลาเสร็จงาน (นาที) ของเว็บไซต์ที่ไม่ได้ทำการปรับเหมาะกับผู้ใช้งาน ใช้เวลานานกว่าการใช้เว็บไซต์ที่ทำการปรับเหมาะกับผู้ใช้งาน โดยมีรายละเอียดผลการทดสอบค่าเฉลี่ยโดยใช้การประเมิน การจับเวลาเสร็จงาน (นาที) สำหรับกลุ่มทดลองทั้ง 4 ประเภท ดังภาพที่ 4

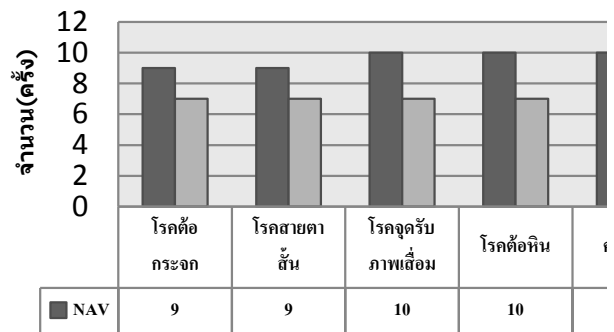


ภาพที่ 4: แผนภาพเปรียบเทียบผลการทดสอบค่าเฉลี่ยโดยใช้ การประเมินการจับเวลาเสร็จงาน (นาที) สำหรับกลุ่มทดลองทั้ง 4 ประเภท

ตารางที่ 3: แสดงการเปรียบเทียบผลการทดสอบโดยใช้การประเมินการนับจำนวนการคลิกระบบนำทาง (ครั้ง) ของเว็บไซต์ที่ไม่ได้ทำการปรับเหมาะกับผู้ใช้งาน และเว็บไซต์ที่ได้ทำการปรับเหมาะกับผู้ใช้งาน

เวอร์ชัน	จำนวน N	ค่าเฉลี่ย \bar{x}	ค่า S.D	ค่า t	Sig.
NAV	40	10	2.18018	9.358	.000
AV	40	7	1.83956		

จากตารางที่ 3 พบว่า การนับจำนวนการคลิกระบบนำทาง (ครั้ง) ของเว็บไซต์ที่ไม่ได้ทำการปรับเหมาะกับผู้ใช้งานมีจำนวนการคลิกมากกว่าโดยมีค่าเฉลี่ยการใช้ 10 ครั้ง ส่วนการใช้เว็บไซต์ที่ทำการปรับเหมาะกับผู้ใช้งานมีค่าเฉลี่ยการคลิกระบบนำทาง 7 ครั้ง โดยมีรายละเอียดผลการทดสอบค่าเฉลี่ยโดยใช้การประเมิน การนับจำนวนการคลิกระบบนำทาง (ครั้ง) สำหรับกลุ่มทดลอง ทั้ง 4 ประเภทดังภาพที่ 5



ภาพที่ 5: แผนภาพเปรียบเทียบผลการทดสอบค่าเฉลี่ยโดยใช้การประเมิน การนับจำนวนระบบนำทาง (ครั้ง) สำหรับกลุ่มทดลองทั้ง 4 ประเภท

4.2 ผลการประเมินโดยใช้แบบสอบถามความคิดเห็นการใช้ประโยชน์จากเว็บ

ตารางที่ 4: ผลการประเมินโดยใช้แบบสอบถามความคิดเห็นการใช้ประโยชน์จากเว็บ

ความคิดเห็นลักษณะการออกแบบ	ค่าเฉลี่ย	
	\bar{x}	S.D
1.คิดว่าจะใช้งานเว็บนี้บ่อยๆ	4.70	0.48
2.เว็บไซต์นี้ใช้งานง่าย	4.40	0.51
3.ต้องการความช่วยเหลือในการใช้งานเว็บไซต์นี้	4.00	0.47
4.มีความเหมาะสมในการใช้สีของตัวอักษรและรูปภาพ	4.00	0.47
5.มีความเหมาะสมในการเลือกใช้ขนาดของตัวอักษรบนจอภาพ	4.20	0.79
6.มีความเหมาะสมในการปฏิสัมพันธ์โต้ตอบกับผู้ใช้	4.50	0.53
7.เว็บไซต์ช่วยให้พบข้อมูลที่ต้องการได้อย่างง่ายดาย	4.30	0.48
8.ค้นพบสิ่งที่ไม่สอดคล้องในการออกแบบ	4.30	0.67
9.เว็บไซต์ใช้งานยากและมีข้อผิดพลาดในการใช้งาน	4.40	0.70
10.รู้สึกมั่นใจในการใช้เว็บไซต์นี้	4.10	0.32
ค่าเฉลี่ย	4.29	0.98

จากตารางที่ 4 พบว่าความเห็นคิดเห็นการใช้ประโยชน์จากเว็บส่วนต่อประสานแบบปรับเหมาะสมกับผู้ใช้ เมื่อพิจารณาความคิดเห็นเป็นรายด้าน ผู้ใช้ให้ความคิดเห็นว่าจะใช้งานเว็บนี้บ่อยๆ พบว่ามีระดับ เห็นด้วยอย่างยิ่งมีค่าเฉลี่ยเท่ากับ 4.70 และผู้ใช้อ้างอิงให้ความคิดเห็นรองลงมาคือมีความเหมาะสมในการปฏิบัติสัมพันธ์ได้ตอบกับผู้ใช้ พบว่ามีระดับ เห็นด้วยอย่างยิ่งมีค่าเฉลี่ย เท่ากับ 4.50 อย่างไรก็ตามกล่าวได้ว่าโดยรวมผู้ใช้งานมีความคิดเห็นผู้ที่มีความพึงพอใจอยู่ในระดับดีโดยมีค่าเฉลี่ยเท่ากับ 4.29

5. สรุปและอภิปรายผล

จากผลการประเมินโดยใช้แบบสอบถามความคิดเห็นการใช้ประโยชน์จากเว็บและการเข้าถึงเนื้อหาโดยการกำหนดงาน (Task) ให้ผู้ทดลองได้ปฏิบัติคำสั่ง ในการเปรียบเทียบผลการทดสอบ โดยใช้การประเมินการนับจำนวนการคลิกกระบวนนำทาง(ครั้ง) ของเว็บไซต์ที่ไม่ได้ทำการปรับเหมาะสมกับผู้ใช้ และเว็บไซต์ที่ได้ทำการปรับเหมาะสมกับผู้ใช้แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 โดยเว็บไซต์ที่ได้ทำการปรับเหมาะสมกับผู้ใช้ มีความสามารถระบบง่ายต่อการใช้งานและการเข้าถึงเนื้อหา มากกว่าเว็บไซต์ที่ไม่ได้ทำการปรับเหมาะสมกับผู้ใช้ ส่วนผลการทดสอบโดยใช้การประเมินการจับเวลาเสร็จงาน(นาท) เว็บไซต์ที่ไม่ได้ทำการปรับเหมาะสมกับผู้ใช้ และเว็บไซต์ที่ได้ทำการปรับเหมาะสมกับผู้ใช้ แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 และพบว่าเว็บไซต์ที่ได้ทำการปรับเหมาะสมกับผู้ใช้ ใช้เวลาในการเข้าถึงเนื้อหาเว็บไซต์ที่ต้องการน้อยกว่า เว็บไซต์ที่ไม่ได้ทำการปรับเหมาะสมกับผู้ใช้ ส่วนความคิดเห็นการใช้ประโยชน์จากเว็บ โดยการใช้แบบสอบถามนั้น โดยรวมผู้ที่มีความพึงพอใจอยู่ในระดับดี

การวิจัยครั้งนี้มีผลตรงตามวัตถุประสงค์ในการศึกษาและตรงตามสมมติฐานที่ตั้งไว้ โดยศึกษาทฤษฎีและวิจัยที่เกี่ยวข้องนำมาวิเคราะห์เป็นแนวทางในศึกษาและพัฒนาการวิจัยครั้งนี้พบว่าผลการวิจัยมีความสอดคล้องกับวิจัยอื่นๆ ในเรื่องประสิทธิภาพส่วนต่อประสานแบบปรับเหมาะสมกับผู้ใช้ในการเข้าถึงเว็บ รวมทั้งปัจจัยต่างๆที่เอื้อต่อการเข้าถึงเว็บสำหรับผู้พิการทางสายตา ประเภทสายตาเลือนราง ประโยชน์ที่ได้รับจากงานวิจัยในครั้งนี้ เป็นแนวคิดในการพัฒนาส่วนต่อประสาน

แบบปรับเหมาะสมสำหรับบุคคลที่มีความบกพร่องประเภทอื่นๆหรือมีความแตกต่างด้านต่างๆเช่น ผู้ที่มีความบกพร่องทางหู หรือความแตกต่างทางอายุ สำหรับการผู้ใช้งานบนเว็ลด์ไวด์เว็บต่อไป

เอกสารอ้างอิง

- [1] Horgen, S. A. 2001 "Essay on Intelligent User Interface and adaptively," *Division of Intelligent Systems*, NTNU.
- [2] Mitchell, J. and Shneiderman, B., "Dynamic versus static menus: an exploratory comparison," *SIGCHI Bull*20, 33-37, 1989.
- [3] Benyon, D, and Murray, D. 1993 "Apply user modeling to human-computer interaction design. *Artificial Intelligence Review*", CERMA 2007.
- [4] Langley, P. 1999, "User Modeling in Adaptive Interface, Proceeding," *Seventh International Conference on User Modeling*. Springer-Verlag, 305-330.
- [5] Talia Lavie & Joachim Meyer. "Benefits and costs of adaptive user interfaces," *Int. J. Human-Computer Studies* 68, 508-524, 2010.
- [6] สำนักงานปลัดกระทรวงเทคโนโลยีสารสนเทศและการสื่อสาร, "แนวทางการพัฒนาเว็บที่ทุกคนเข้าถึงได้" สืบค้นเมื่อ 30 สิงหาคม 2556 จาก lib02.kku.ac.th/kb_upload/File/doc/TWCAG2010.doc.
- [7] จุฬารัตน์ แซ่เนี่ยม, "การปรับปรุงการเข้าใช้งานเว็บสำหรับผู้พิการทางสายตา," วิทยานิพนธ์ปริญญาโทมหาบัณฑิต. ภาควิชาการคอมพิวเตอร์ บัณฑิตวิทยาลัย มหาวิทยาลัยมหิดล, 2007.
- [8] "Types of Low Vision" [Online] [.http://www.http://webaim.org](http://www.http://webaim.org), 28, August 2013. lib02.kku.ac.th/kb_upload/File/doc/TWCAG2010.doc.
- [9] Gloria, A, Reece, "Usable, Accessible Web Page for Low Vision: Criteria for Designers," *Instructional Design and Technology School of Education The University of Memphis*.
- [10] Marguerite Bergel, Ann Chadwick-Dias, Lori LeDoux, and Tom Tullis, "Web Accessibility for the Low Vision User," *Usability Professionals Association (UPA) 2005 Presentation*, Montreal, Quebec, Canada.

การพัฒนาเว็บไซต์เพื่อสุขภาพตามแนวทาง TWCAG 2010 สำหรับผู้สูงอายุ

A Development of Health-related Websites for Aging Users Based on TWCAG 2010

ดวงใจ หนูเล็ก (Duangjai Noolek)¹ และทิพย์ยา จินตโกวิท (Thippaya Chintakovid)²

^{1,2}ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹ duangjai.noolek@hotmail.com, ² thippayac@kmutnb.ac.th

บทคัดย่อ

เว็บไซต์ด้านสุขภาพสำหรับผู้สูงอายุไทยในปัจจุบันมีจำนวนไม่มากและยากต่อการทำความเข้าใจข้อมูลของผู้สูงอายุ การทดสอบการเข้าถึงเว็บไซต์สำหรับผู้สูงอายุด้วยเครื่องมือตรวจสอบจาก ThaiWebAccessibility.com พบว่าเว็บไซต์สำหรับผู้สูงอายุส่วนใหญ่ไม่ได้พัฒนาตามแนวทางการออกแบบเว็บไซต์ที่ช่วยให้ทุกคนสามารถเข้าถึงได้ งานวิจัยนี้จึงพัฒนาเว็บไซต์ด้านสุขภาพสำหรับผู้สูงอายุ ตามแนวทางการพัฒนาเว็บไซต์ที่ทุกคนเข้าถึงได้ฉบับภาษาไทย (TWCAG 2010) เพื่อช่วยให้ผู้สูงอายุเข้าถึงและใช้งานเนื้อหาบนเว็บไซต์ได้ง่ายขึ้น ซึ่งเว็บไซต์ที่พัฒนาขึ้นผ่านการทดสอบการเข้าถึงเว็บไซต์ที่ระดับสูงสุด จากการประเมินความยากง่ายในการเรียนรู้การใช้งานเว็บไซต์และความพึงพอใจโดยกลุ่มตัวอย่างผู้สูงอายุจำนวน 30 คน พบว่าผู้สูงอายุเรียนรู้ในการใช้งานเว็บไซต์ที่พัฒนาขึ้นง่ายกว่า และมีความพึงพอใจต่อเว็บไซต์ที่พัฒนาขึ้นมากกว่า

คำสำคัญ: แนวทางการพัฒนาเว็บไซต์ที่ทุกคนเข้าถึงได้ฉบับภาษาไทย แนวทางการพัฒนาเว็บไซต์เพื่อผู้สูงอายุ เว็บไซต์ด้านสุขภาพ ผู้สูงอายุ

Abstract

There are not many health-related websites available today for Thai elders. Moreover, it is difficult for elders to understand the information provided on these websites. By using a tool from ThaiWebAccessibility.com, a web accessibility test showed that most of today's websites for elders do not follow the web accessibility guideline. This research developed a health related websites for aging users based on Thai Web Content Accessibility Guideline

2010 (TWCAG 2010) to help old users to more easily access and use information on the web. The developed website was tested and passed the highest accessibility level of the guideline. The learnability and satisfaction of the website was evaluated by 30 old users. Results revealed that it was easier for the users to learn to use and more satisfactory to use the developed website than an existing website.

Keywords: Thai Web Content Accessibility Guideline 2010, Web Design Guidelines for Older People, Health related Websites, Aging Users.

1. บทนำ

ประเทศไทยได้เข้าสู่สังคมผู้สูงอายุตั้งแต่ปี พ.ศ. 2550 [1] และเทคโนโลยีสารสนเทศเป็นสิ่งที่มีความประโยชน์ที่จะช่วยผู้สูงอายุในการทำงานหรือประกอบอาชีพหลังเกษียณอายุงานอย่างสะดวกมากขึ้น [2] และเป็นอีกช่องทางหนึ่งที่ช่วยให้ผู้สูงอายุค้นหาหาข้อมูลที่นอกเหนือจากการทำงานตามที่ต้องการ

จากงานวิจัยที่เกี่ยวกับพฤติกรรมกรรมการแสวงหาสารสนเทศของผู้สูงอายุพบว่า ผู้สูงอายุค้นหาข้อมูลด้านสุขภาพมากที่สุด [3], [4] แต่เว็บไซต์ด้านสุขภาพสำหรับผู้สูงอายุไทยมีจำนวนไม่มาก และเมื่อทดสอบเว็บไซต์ดังกล่าวด้วยเครื่องมือตรวจสอบจาก ThaiWebAccessibility.com พบว่าเว็บไซต์ส่วนใหญ่ไม่ได้พัฒนาตามแนวทางในการออกแบบเว็บไซต์ที่ช่วยให้ทุกคนสามารถเข้าถึงได้ นอกจากนี้ ผู้สูงอายุส่วนใหญ่มีความเห็นว่าการใช้งานเทคโนโลยีสารสนเทศ มีความยุ่งยากซับซ้อน และสารสนเทศบนอินเทอร์เน็ตมีมากเกินไป หรือข้อมูลที่ค้นหาได้ไม่ตรงตามความต้องการ [3] สำหรับใช้เป็น

ข้อมูลในการทำงานหรือดูแลสุขภาพของตนเองได้ จากปัญหาดังกล่าว ผู้วิจัยจึงพัฒนาเว็บไซต์เกี่ยวกับสุขภาพที่เหมาะสมกับผู้สูงอายุ ตามแนวทางการพัฒนาเว็บไซต์ที่ทุกคนเข้าถึงได้ฉบับภาษาไทย (Thai Web Content Accessibility Guideline 2010 : TWACAG 2010) [5], [6], [7] ที่จัดทำขึ้นโดยสำนักปลัดกระทรวงเทคโนโลยีสารสนเทศและการสื่อสาร ซึ่งอ้างอิงจากมาตรฐานสากล WCAG 2.0 [8], [9] และพัฒนาเว็บไซต์ตามแนวทางการออกแบบเว็บไซต์สำหรับผู้สูงอายุจากงานวิจัยที่ผ่านมา [10], [11] เพื่อช่วยให้ผู้สูงอายุสามารถเข้าถึงการใช้งานเว็บไซต์ได้ง่าย และสามารถค้นหาข้อมูลบนเว็บไซต์เพื่อการดูแลรักษาสุขภาพของตนเองในปัจจุบันได้

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 แนวทางการพัฒนาเว็บไซต์สำหรับผู้สูงอายุ

วัยสูงอายุเป็นวัยที่มีปัญหาสุขภาพที่เสื่อมถอย ทำให้ความสามารถต่าง ๆ เช่น ความคิด การกระทำ และการเรียนรู้ช้าลง การมองเห็นและการได้ยินลดลง [2], [3] ทำให้ประสบความยากลำบากในการใช้งานเทคโนโลยีคอมพิวเตอร์และอินเทอร์เน็ต จึงมีงานวิจัยที่ศึกษาและพัฒนาแนวทางในการพัฒนาเว็บไซต์ [10], [11] ที่จะช่วยเพิ่มความสามารถในการเข้าถึงและใช้งานเว็บไซต์ของผู้สูงอายุ สามารถสรุปรายละเอียดของแนวทางต่างๆ ดังนี้

2.1.1 แนวทางสำหรับการออกแบบข้อความ ใช้แบบอักษรที่ไม่เป็นตัวหนาเพื่อแสดงเนื้อหาข้อมูล หลีกเลี่ยงใช้สีที่ใกล้เคียงกัน เช่น สีเหลือง สีฟ้า และสีเขียว และควรใช้สีเข้มบนพื้นหลัง เป็นต้น

2.1.2 แนวทางสำหรับการนำเสนอข้อมูล นำเสนอข้อมูลในรูปแบบที่ชัดเจน ใช้ภาษาที่เข้าใจง่าย จัดระเบียบเนื้อหาในรูปแบบมาตรฐาน

2.1.3 แนวทางสำหรับการเพิ่มความง่ายโดยใช้ เนวิกชัน (navigation) ออกแบบหน้าจอแบบมาตรฐานและใช้สัญลักษณ์ไอคอนที่เหมือนกันตลอดทุกหน้าจอและอยู่ในตำแหน่งเดียวกัน มีระบบนำทางและมีเมนูช่วยเหลือสำหรับผู้ที่ต้องการสอบถามหรือแสดงความคิดเห็น

2.2 การพัฒนาเว็บไซต์ที่ทุกคนเข้าถึงได้ (Web accessibility)

ความสามารถในการเข้าถึงได้ทางเว็บไซต์ หมายถึง การ

พัฒนาเว็บไซต์ที่รองรับผู้ใช้ทุกกลุ่ม ไม่ว่าจะเป็นผู้พิการหรือบุคคลทั่วไป การที่ทำให้เว็บไซต์มีความสามารถเช่นนี้ จะทำให้ผู้ใช้สามารถเข้าถึงและเข้าใจเนื้อหา (Content) ในเว็บไซต์นั้นได้ และเป็นประโยชน์กับคนพิการ ผู้สูงอายุที่สายตาเลือนลาง จนกระทั่งผู้ใช้ปกติที่ใช้ช่องสัญญาณความเร็วต่ำ ให้สามารถเข้าถึงเนื้อหาในเว็บไซต์ได้อย่างเท่าเทียมกัน [12], [13]

สำหรับประเทศไทย กระทรวงเทคโนโลยีสารสนเทศและการสื่อสาร (ICT) ได้จัดทำแนวทาง TWACAG ซึ่งเป็นแนวทางการพัฒนาเว็บไซต์ที่ทุกคนสามารถใช้งานและเข้าใจเนื้อหาได้ โดยมีหลักการดังนี้

หลักการที่ 1 ผู้อ่านต้องสามารถรับรู้เนื้อหาได้ โดยการเห็นได้ยินหรือสัมผัส สามารถปฏิบัติได้โดยจัดให้มีข้อความทดแทนสำหรับเนื้อหาที่ไม่ใช่ข้อความหรือสื่อประเภทเสียงหรือวีดิทัศน์ที่มีการบันทึกไว้ล่วงหน้า เพื่อเพิ่มความสามารถในการรับรู้เนื้อหาของผู้ใช้ผ่านข้อความที่เป็นคำอธิบาย ออกแบบเว็บไซต์ให้ผู้ใช้สามารถปรับเปลี่ยนรูปแบบได้ตามความเหมาะสมและตามความต้องการ โดยไม่สูญเสียเนื้อหาหรือการทำงานเดิม เช่น การปรับขนาดตัวอักษร การหยุดหรือการปิดเสียงประกอบภาพ หากมีการใส่เสียงประกอบที่มีความยาวมากกว่า 3 วินาที ต้องสามารถหยุดเล่นชั่วคราวหรือจบการเล่นได้ เป็นต้น

หลักการที่ 2 องค์ประกอบต่างๆ ของหน้าจอกับเนื้อหาต้องใช้งานได้ ผู้ใช้สามารถเข้าถึงได้ทุกส่วนและใช้งานในหน้าเว็บไซต์ได้ด้วยการใช้แป้นพิมพ์ ซึ่งจะทำให้ผู้ใช้ที่มีข้อจำกัดด้านฮาร์ดแวร์ หรือผู้พิการทางการมองเห็นสามารถเข้าถึงเนื้อหาได้ รวมถึงการใช้แอนิเมชันที่เหมาะสมกับการรับรู้เนื้อหา โดยไม่ส่งผลกระทบต่อผู้ที่มีความบกพร่องทางร่างกาย เช่น การสร้างเนื้อหาบนเว็บไซต์ที่มีลักษณะแสงกระพริบเกิน 3 ครั้งต่อ 1 วินาที จะมีผลกระทบต่อผู้ใช้ที่เป็นโรคลมชักซึ่งจะมีอาการไวต่อการกระตุ้นทางแสง อาจทำให้เกิดอาการโรคลมชักกำเริบขึ้นมาได้ เป็นต้น

หลักการที่ 3 ผู้ใช้สามารถเข้าใจเนื้อหา ส่วนควบคุมการทำงานต่างๆ และส่วนต่อประสาน โดยจัดวางโครงสร้างเนื้อหาให้เป็นระเบียบ และมีรูปแบบที่เหมือนกันทุกหน้าอย่างเหมาะสม เพื่อให้ผู้ใช้เกิดความคุ้นเคยและสามารถคาดเดาเนื้อหาของเว็บไซต์ที่ต้องการได้ว่าอยู่ส่วนใดของหน้าเว็บ ควรมีคำแนะนำการใช้งาน

ในบางฟังก์ชันที่สำคัญ เช่น การกรอกข้อมูลในแบบฟอร์ม เมื่อผู้ใช้อกรอกข้อมูลผิดประเภทควรมีข้อความแจ้งเตือนเพื่อให้ผู้ใช้เกิดความเข้าใจและ หลีกเลี่ยงความผิดพลาดที่จะเกิดขึ้นซ้ำอีก เป็นต้น

หลักการที่ 4 เนื้อหาต้องมีความยืดหยุ่นเหมาะที่จะทำงานกับเทคโนโลยีเว็บไซต์ในปัจจุบันและอนาคต โดยออกแบบเว็บไซต์ที่มีการสร้างเนื้อหาให้รองรับการใช้งานกับเทคโนโลยีที่หลากหลาย แม้ว่าการพัฒนาของเทคโนโลยีจะเปลี่ยนไป เช่น โปรแกรมบราวเซอร์ และอุปกรณ์ต่าง ๆ [5], [6], [13]

ในการพัฒนาเว็บไซต์ที่ทุกคนเข้าถึงได้ มีเครื่องมืออัตโนมัติในการตรวจสอบว่าเว็บไซต์ได้รับการพัฒนาตาม TWCAG 2010 อยู่ในระดับใด [14] โดยกำหนดระดับในการเข้าถึงไว้ 3 ระดับคือระดับ A ระดับ AA ระดับ AAA ซึ่งระดับ AAA เป็นระดับที่ช่วยให้ผู้ใช้งานเข้าถึงเว็บไซต์ได้ง่ายที่สุด [5], [6], [13]

3. วิธีการดำเนินการวิจัย

3.1 สํารวจเว็บไซต์ที่เกี่ยวข้องกับผู้สูงอายุ

ผู้วิจัยสำรวจเว็บไซต์สำหรับผู้สูงอายุ โดยค้นหาเว็บไซต์สำหรับผู้สูงอายุในปัจจุบันบน www.google.com และคัดเลือกเว็บไซต์ที่เป็นที่นิยมจำนวน 10 เว็บไซต์ โดยพิจารณาความนิยมจากจำนวนผู้เข้าชมเว็บไซต์ จากนั้นทำการทดสอบแต่ละเว็บไซต์ด้วยเครื่องมือตรวจสอบจาก ThaiWeb Accessibility.com ผลการตรวจสอบเว็บไซต์ทั้งหมด พบว่าร้อยละ 80 ของเว็บไซต์ ไม่ผ่านเกณฑ์ตามแนวทาง TWCAG 2010 ดังแสดงในตารางที่ 3-1

ตารางที่ 3-1: ผลการตรวจสอบเว็บไซต์ด้วยเครื่องมือตรวจสอบจาก ThaiWebAccessibility.com

ชื่อเว็บไซต์	ผลการตรวจสอบ	
	ผ่าน	ไม่ผ่าน
1. www.oldisyounginbkk.com	√	
2. http://www.happyoppy.com/		√
3. https://www.gpf.or.th		√
4. http://www.oppo.opp.go.th/		√
5. http://www.opp.go.th/		√
6. http://www.thaicentarian.mahidol.ac.th/		√
7. http://www.thehealthyclub.com/	√	
8. http://www.goodfoodgoodlife.in.th/		√
9. www.ผู้สูงอายุ.net		√
10. http://www.banbangkhae.go.th/		√

3.2 ออกแบบและพัฒนาเว็บไซต์

3.2.1 ผู้วิจัยออกแบบเว็บไซต์ตามแนวทางการพัฒนาเว็บไซต์ที่ทุกคนเข้าถึงได้ฉบับภาษาไทย ดังตารางที่ 3-2

ตารางที่ 3-2: เกณฑ์การออกแบบการเข้าถึงขั้นสูงสุดที่ช่วยให้ผู้ใช้เข้าถึงและใช้งานเนื้อหาเว็บไซต์ได้ง่ายขึ้น

หลักเกณฑ์	การออกแบบ
ไม่สร้างเสียงพื้นหลัง	ไม่ใส่เสียงพื้นหลัง เลือกภาพพื้นหลังเป็นสีอ่อน
ขนาดของตัวอักษรสามารถปรับขนาดได้ ร้อยละ 50-200 โดยปราศจากเครื่องมือช่วยอื่น และไม่ทำให้ความหมายในเนื้อหาเปลี่ยนแปลง และต้องไม่ทำให้เกิดการ Scroll ทางแนวนอน	ผู้ใช้งานสามารถเลือกปรับขนาดตัวอักษรได้เอง 3 ระดับ และ หน้าเว็บ ไม่มี Scroll แนวนอน
เนื้อหาจะต้องไม่มีข้อมูลที่เป็นลักษณะแฟลช ที่มีการกะพริบมากกว่า 3 ครั้งต่อวินาที	ส่วนของเนื้อหาไม่ใส่ข้อมูลที่เป็นลักษณะแฟลช
มีส่วนที่แสดงให้ทราบถึงเส้นทางตำแหน่งของหน้าเว็บ	มีเมนูแผนผังเว็บไซต์ และระบบเนวิเกชัน
ชื่อลิงค์บอกความหมายในตัวเองได้	ชื่อของลิงค์ในแต่ละหน้าเว็บบอกความหมายชัดเจน
การแบ่งหมวดหมู่ให้ใช้ Heading แบ่งเป็นลำดับชั้น มีรูปแบบที่ชัดเจน จักระเบียบเนื้อหาในรูปแบบมาตรฐาน	จัดรูปแบบ Heading เป็นตัวอักษรตัวหนาสีดำ
จัดเตรียมระบบการช่วยเหลือสำหรับวิธีการกรอกข้อมูลที่ถูกต้อง	หากผู้ใช้งานมีการกรอกข้อมูลผิดประเภทจะ มีข้อความแจ้งเตือนและคำแนะนำว่าให้กรอกข้อมูลประเภทใด
มีวิธีป้องกันการเกิดความผิดพลาดอย่างน้อย 1 วิธี	ผู้ใช้งานที่เป็นสมาชิกสามารถกลับไปแก้ไขหรือตรวจสอบข้อมูลของสมาชิกเองได้
โครงสร้างเอกสารต้องเป็นไปตามมาตรฐานของภาษา	ออกแบบตามหลักการของภาษา HTML

3.2.2 ผู้วิจัยออกแบบเว็บไซต์ตามหลักการออกแบบเว็บไซต์เพื่อผู้สูงอายุ [10], [11] ดังตารางที่ 3-3 และพัฒนาขึ้นในรูปแบบของเว็บแอปพลิเคชัน ด้วยภาษา PHP ร่วมกับโปรแกรมจัดการฐานข้อมูล MySQL ใช้เทคโนโลยี CSS (Cascading Style Sheets) พัฒนาหน้าเว็บไซต์ให้มีความเป็นระเบียบและสวยงามมากขึ้น

ตารางที่ 3-3: หลักการออกแบบเว็บไซต์สำหรับผู้สูงอายุ

แนวทางสำหรับการออกแบบข้อความ	
แบบอักษร	ใช้แบบอักษร Adamina ตัวอักษรธรรมดา สีดำบนพื้นสีอ่อน และสีขาวบนพื้นสีเข้ม
แนวทางสำหรับการนำเสนอข้อมูล	
รูปแบบ	มีการจัดรูปแบบที่เหมือนกันทุกหน้า
ความง่าย	เขียนข้อความภาษาไทยที่อ่านง่าย ใช้คำศัพท์ทั่วไป
แนวทางสำหรับการเพิ่มความง่ายในการใช้เนวิเกชัน (navigation)	
ช่วยเหลือและติดต่อ	มีฟังก์ชันค้นหาสำหรับผู้ใช้ที่ต้องการค้นหาข้อมูล และมีเมนูติดต่อ สำหรับผู้ใช้ที่ต้องการติดต่อกับผู้ดูแลเว็บไซต์
Site Map	มีเมนูแผนผังเว็บไซต์
เมนู	ออกแบบเมนูอยู่ด้านบนเหมือนกันทุกหน้าเพื่อง่ายต่อการมองเห็นและใช้งาน

3.3 การทดสอบเว็บไซต์

ผู้วิจัยทดสอบการทำงานของเว็บไซต์ด้วยเครื่องมือตรวจสอบจาก ThaiWebAccessibility.com ในขั้นสูงสุดที่ช่วยให้ผู้ใช้เข้าถึงและใช้งานเนื้อหาเว็บได้ง่ายขึ้น (AAA) และประเมินความยากง่ายในการเรียนรู้การใช้งานและความพึงพอใจจากกลุ่มตัวอย่างอายุ 50 ปีขึ้นไป ที่สามารถใช้งานคอมพิวเตอร์และอินเทอร์เน็ตได้จำนวน 30 คน [15] ขนาดของกลุ่มตัวอย่างคำนวณด้วยโปรแกรม G*POWER โดยเลือกวิธีการทางสถิติแบบ paired sample t-test และกำหนด Effect size = 0.5, $\alpha = .05$ และ Power = 0.8 งานวิจัยนี้คัดเลือกกลุ่มตัวอย่างที่มีอายุ 50 ปีขึ้นไป เนื่องจากผลการสำรวจพฤติกรรมผู้ใช้อินเทอร์เน็ตในประเทศไทย ปี 2556 พบว่าผู้ใช้อินเทอร์เน็ตอายุ 50-59 ปีมีพฤติกรรมการใช้งานอินเทอร์เน็ตคล้ายคลึงกับผู้ใช้อินเทอร์เน็ตอายุ 60 ปีขึ้นไป และจำนวนผู้ตอบแบบสอบถามที่มีอายุระหว่าง 50-59 ปีมีมากกว่าผู้ตอบแบบสอบถามที่มีอายุ 60 ปีขึ้นไป [15] ในส่วนกิจกรรมเพื่อทดสอบการใช้งานเว็บไซต์ ผู้วิจัยกำหนดกิจกรรม 5 อย่าง ให้กลุ่มตัวอย่างทดสอบการใช้งานเว็บไซต์ที่พัฒนาขึ้น (www.healthyforaging.com) เปรียบเทียบกับการใช้งานเว็บไซต์ด้านสุขภาพสำหรับผู้สูงอายุที่มีอยู่ในปัจจุบัน (www.thehealthyclub.com) กิจกรรมดังกล่าว ได้แก่ สมัครสมาชิกเว็บไซต์ อ่านเมนูอาหารด้านสุขภาพที่สนใจอ่านบทความที่สนใจคำนวณค่าพลังงานของอาหาร และติดต่อผู้ดูแลเว็บไซต์

ซึ่งเป็นกิจกรรมที่เกี่ยวข้องกับฟังก์ชันพื้นฐานที่มีในเว็บไซต์ทั้ง 2 เว็บไซต์ กลุ่มตัวอย่างให้คะแนนความยากง่ายในการเรียนรู้การใช้งานเว็บไซต์จากคะแนน 1 (ง่ายที่สุด) ถึง 5 (ยากที่สุด) และให้คะแนนความพึงพอใจต่อประสิทธิภาพและการออกแบบเว็บไซต์จากคะแนน 1 (น้อยที่สุด) ถึง 5 (มากที่สุด)

4. ผลการดำเนินงาน

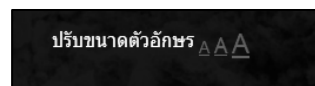
เว็บไซต์ถูกพัฒนาขึ้นในรูปแบบของเว็บแอปพลิเคชัน ด้วยภาษา PHP ร่วมกับโปรแกรมจัดการฐานข้อมูล MySQL ใช้เทคโนโลยี css (Cascading Style Sheets) พัฒนาหน้าเว็บไซต์ให้มีความเป็นระเบียบและสวยงามมากขึ้น ตามหลักการออกแบบเว็บไซต์สำหรับผู้สูงอายุ และตามแนวทางWCAG 2010 โดยมีผลการดำเนินงานดังนี้

4.1 ออกแบบสีพื้นหลังเป็นสีอ่อนและไม่ใส่เสียง ชื่อของลิงค์ในแต่ละหน้าเว็บเพจออกความหมายชัดเจนและอยู่ด้านบนเหมือนกันทุกหน้าเพื่อง่ายต่อการมองเห็นและใช้งาน ส่วนเนื้อหาไม่มีข้อมูลที่เป็นลักษณะแฟลช มีการแสดงตัวอักษรสีต่างจากสีปกติเมื่อนำเมาส์ชี้ หรือในขณะที่ผู้ใช้งานอยู่ในหน้าที่เลือก ดังแสดงในภาพที่ 4-1



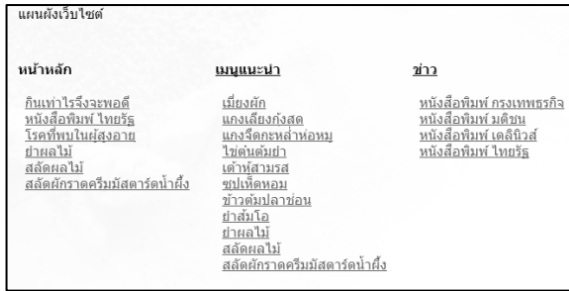
ภาพที่ 4-1: แสดงสีพื้นหลังเป็นสีอ่อนและไม่ใส่เสียง

4.2 สามารถปรับขนาดตัวอักษรได้ 3 ระดับโดยที่ไม่ทำให้เนื้อหาในหน้าเว็บไซต์เปลี่ยนแปลง ดังแสดงในภาพที่ 4-2



ภาพที่ 4-2: แสดงการปรับขนาดตัวอักษรได้ 3 ระดับ

4.3 แผนผังเว็บไซต์แสดงให้เห็นทราบถึงเส้นทางตำแหน่งของเว็บเพจดังแสดงในภาพที่ 4-3



ภาพที่ 4-3: แสดงแผนผังเว็บไซต์

4.4 มีฟังก์ชันค้นหาสำหรับผู้ใช้งานที่ต้องการค้นหาข้อมูล ดังแสดงในภาพที่ 4-4



ภาพที่ 4-4: แสดงฟังก์ชันค้นหาข้อมูล

4.5 มีข้อความแจ้งเตือนผู้ใช้งานในการกรอกข้อมูล แนะนำให้กรอกข้อมูลประเภทใด ดังแสดงในภาพที่ 4-5



ภาพที่ 4-5: แสดงข้อความแจ้งเตือน

4.6 ออกแบบให้สามารถเปิดใช้งานเว็บไซต์ได้ทั้งในเครื่องคอมพิวเตอร์ โน้ตบุ๊ก ไอแพด หรือมือถือ และยังเปิดได้ในโปรแกรมบราวเซอร์ต่างๆ ดังแสดงในภาพที่ 4-6



ภาพที่ 4-6: แสดงการเปิดใช้งานเว็บไซต์บนมือถือ

4.7 ผลการทดสอบการทำงานของเว็บไซต์ตามแนวทาง TWAG2010 ที่ระดับ AAA ด้วยเครื่องมือตรวจสอบจาก ThaiWebAccessibility.com พบว่าเว็บไซต์ที่พัฒนาขึ้นผ่านการทดสอบแบบมีเงื่อนไข ดังแสดงในภาพที่ 4-7 เนื่องจากเครื่องมือตรวจสอบไม่สามารถประมวลผลข้อความภาษาไทยที่

เป็นค่าฟิลด์ของแท็ก HTML ได้ จึงทำให้เว็บไซต์ไม่ผ่านเกณฑ์การตรวจสอบบางส่วน เช่น เครื่องมือตรวจสอบไม่สามารถเข้าใจข้อความบรรยายภาพภาษาไทยที่ระบุในแท็ก alt จึงแจ้งเตือนว่าข้อความในแท็ก alt ไม่สื่อความหมายเช่นเดียวกับภาพ เป็นต้น



ภาพที่ 4-7: แสดงผลการทดสอบเว็บไซต์ ที่ระดับ AAA ด้วยเครื่องมือตรวจสอบจาก ThaiWebAccessibility.com

4.8 ผลการประเมินความยากง่ายในการเรียนรู้และความพึงพอใจในการใช้งานเว็บไซต์ กลุ่มผู้สูงอายุที่ประเมินการใช้งานเว็บไซต์เป็นเพศชายจำนวน 20 คน (ร้อยละ 66.67) และเพศหญิงจำนวน 10 คน (ร้อยละ 33.33) มีอายุระหว่าง 50-59 ปี จำนวน 21 คน (ร้อยละ 70) และอายุระหว่าง 60-69 ปี จำนวน 9 คน (ร้อยละ 30) สำหรับอาชีพของกลุ่มตัวอย่าง มีอาชีพ รับราชการ จำนวน 14 คน (ร้อยละ 46.67) เป็นพนักงานบริษัทจำนวน 2 คน (ร้อยละ 6.67) ประกอบธุรกิจส่วนตัวจำนวน 6 คน (ร้อยละ 20) และเป็นข้าราชการบำนาญจำนวน 8 คน (ร้อยละ 26.67) กลุ่มตัวอย่างส่วนใหญ่มีประสบการณ์ในการใช้งานคอมพิวเตอร์และอินเทอร์เน็ต โดยมีประสบการณ์น้อยกว่า 1 ปีจำนวน 1 คน (ร้อยละ 3.33) มีประสบการณ์ระหว่าง 1-3 ปี จำนวน 6 คน (ร้อยละ 20) ระหว่าง 4-6 ปีจำนวน 9 คน (ร้อยละ 30) ระหว่าง 7-9 ปีจำนวน 4 คน (ร้อยละ 13.33) และ 10 ปีขึ้นไปจำนวน 10 คน (ร้อยละ 33.33) ส่วนความถี่ในการใช้งานอินเทอร์เน็ต มีการใช้งานทุกวันจำนวน 23 คน (ร้อยละ 76.67) ใช้งาน 2-3 ครั้ง/สัปดาห์จำนวน 6 คน (ร้อยละ 20) และใช้งานนาน ๆ ครั้งจำนวน 1 คน (ร้อยละ 3.33)

ผลการเปรียบเทียบคะแนนความยากง่ายในการเรียนรู้การใช้งานและความพึงพอใจระหว่างเว็บไซต์ที่พัฒนาขึ้นกับเว็บไซต์ที่มีในปัจจุบันด้วย Paired Sample t-Test พบว่าการเรียนรู้การใช้งาน www.healthyforaging.com (ค่าเฉลี่ยเท่ากับ 1.78) ง่ายกว่า www.thehealthyclub.com (ค่าเฉลี่ยเท่ากับ

2.68) อย่างมีนัยสำคัญทางสถิติที่ 0.05 นอกจากนี้ ผลการประเมินพบว่าผู้สูงอายุมีความพึงพอใจต่อเว็บไซต์ของงานวิจัยนี้ (ค่าเฉลี่ยเท่ากับ 4.29) มากกว่าเว็บไซต์ที่มีอยู่เดิม (ค่าเฉลี่ยเท่ากับ 3.57) ที่ระดับนัยสำคัญทางสถิติที่ 0.05

5. สรุปผลและข้อเสนอแนะ

งานวิจัยนี้ได้ออกแบบและพัฒนาเว็บไซต์เพื่อสุขภาพสำหรับผู้สูงอายุตามแนวทาง TWCAG 2010 ซึ่งมีขั้นตอนในการดำเนินงาน โดยเริ่มจากการสำรวจเว็บไซต์ที่เป็นที่นิยมของผู้สูงอายุ จากนั้นทดสอบแต่ละเว็บไซต์ว่าเป็นไปตามแนวทางการพัฒนาเว็บไซต์ให้ทุกคนเข้าถึงได้หรือไม่ ขั้นตอนต่อมาคือ ทำการออกแบบและพัฒนาเว็บไซต์ตามแนวทางการเข้าถึงเนื้อหาบนเว็บไซต์ฉบับภาษาไทย และขั้นตอนสุดท้ายคือการทดสอบและประเมินเว็บไซต์

งานวิจัยนี้ออกแบบเว็บไซต์ตามระดับการเข้าถึงขั้นสูงสุดที่ช่วยให้ผู้ใช้เข้าถึงและใช้งานเนื้อหาเว็บได้ง่าย โดยไม่ใส่เสียงพื้นหลัง ใช้ภาพพื้นหลังที่มีสีอ่อน ผู้ใช้สามารถปรับขนาดตัวอักษรได้ ไม่ใช้การเลื่อนหน้าเว็บ (Scroll) ตามแนวนอน ไม่ใส่ข้อมูลแฟลช (Flash) มีแผนผังเว็บไซต์ มีชื่อของลิงก์ในแต่ละหน้าเว็บเพจที่สื่อความหมายชัดเจน จัดรูปแบบ Heading เป็นตัวอักษรตัวหนา สีดำ และมีข้อความแจ้งเตือนแนะนำว่าให้กรอกข้อมูลประเภทใด

ผลการประเมินความยากง่ายในการเรียนรู้การใช้งานและความพึงพอใจต่อการใช้งานเว็บไซต์ พบว่าเว็บไซต์ที่พัฒนาขึ้นสามารถเรียนรู้การใช้งานได้ง่ายโดยมีค่าเฉลี่ยเท่ากับ 1.78 และผู้สูงอายุมีความพึงพอใจต่อการใช้งานเว็บไซต์ในระดับมากโดยมีค่าเฉลี่ยเท่ากับ 4.29

ในอนาคต ส่วนของการปรับเปลี่ยนการแสดงผล ควรพัฒนาให้ผู้ใช้สามารถเลือกเปลี่ยนสีพื้นหลัง และเปลี่ยนสีตัวอักษรเป็นแบบสีเข้มหรือสีอ่อนได้ เพื่อรองรับการใช้งานของผู้ใช้งานที่ตาบอดสี และเป็นต้นแบบสำหรับการพัฒนาเว็บไซต์ที่เหมาะสมกับผู้สูงอายุไทยต่อไป

เอกสารอ้างอิง

- [1] เกื้อ วงศ์บุญสิน, "ประชากรศาสตร์ทรัพยากรมนุษย์," จุฬาลงกรณ์มหาวิทยาลัย กรุงเทพฯ, 2547.
- [2] สมาน ลอยฟ้า, "ผู้สูงอายุกับเทคโนโลยีสารสนเทศ," *วารสารสารสนเทศศาสตร์* ปีที่ 29 ฉบับที่ 2 พ.ค.-ส.ค. 2554 หน้า 54-64.

- [3] ชิติพร ขานุกสิริวัฒน์, "พฤติกรรมกรรมการแสวงหาสารสนเทศของผู้สูงอายุ," *วารสารสารสนเทศศาสตร์* ปีที่ 30 ฉบับที่ 2 พ.ค.-ส.ค. 2555 หน้า 77-90.
- [4] ฝนทิพย์ กำลั้งยิ่ง และนชกฤต วันตะเฒ่า, "พฤติกรรมกรรมการแสวงหาสารสนเทศเพื่อวางแผนหลังเกษียณอายุของวัยทำงานตอนปลายในกรุงเทพมหานคร," *การประชุมวิชาการแห่งชาติ ครั้งที่ 9 มหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตกำแพงแสน ธันวาคม 2555* หน้า 850-859.
- [5] สำนักงานปลัดกระทรวงเทคโนโลยีสารสนเทศและการสื่อสาร กระทรวงเทคโนโลยีสารสนเทศและการสื่อสาร, "แนวทางการพัฒนาเว็บที่ทุกคนเข้าถึงได้," กรุงเทพฯ, 2553.
- [6] สำนักงานปลัดกระทรวงเทคโนโลยีสารสนเทศและการสื่อสาร กระทรวงเทคโนโลยีสารสนเทศและการสื่อสาร, "รูปแบบการพัฒนาเว็บที่ทุกคนเข้าถึงได้," กรุงเทพฯ, 2553.
- [7] สำนักงานปลัดกระทรวงเทคโนโลยีสารสนเทศและการสื่อสาร กระทรวงเทคโนโลยีสารสนเทศและการสื่อสาร, "หลักสูตรการพัฒนาเว็บที่ทุกคนเข้าถึงได้ด้วย HTML5 และ CSS ระดับเบื้องต้นและระดับกลาง," กรุงเทพฯ, 2553.
- [8] M. Eyadat and D. Fisher, "Web accessibility in Information Systems," *International Journal of Web*, Vol. 3, No. 4, pp.363-377, 2009.
- [9] T. Miquel, R. Mireia, P. Merce, B. Marc, S. Andreu, and P. Pilar, "Web Content Accessibility Guidelines: from 1.0 to 2.0," *International World Wide Web Conference(www2009)*, pp.1171-1172, 2009.
- [10] A. Shirley, "A Study of Web Usability for Older Adults Seeking Online Health Resources," *ACM Transactions on Computer-Human Interaction*, Vol. 11, no.4, pp. 387-406, 2004.
- [11] K. Sri, and Z. Panayiotis, "Research-Derived Web Design Guidelines for Older People," *ACM Transactions on Computer-Human Interaction*, Vol. 5, pp. 129-135, 2005.
- [12] ปรีดี ปลื้มสำราญกิจ, "ความสามารถในการเข้าถึงได้ทางเว็บ," *วารสารบรรณศาสตร์* ปีที่ 4 ฉบับที่ 2 ก.ค.-ธ.ค. 2554 หน้า 96-106.
- [13] ไฉน ผึ้งผาย, "การพัฒนาเว็บที่ทุกคนสามารถเข้าถึงได้กรณีศึกษา : เว็บไซต์โครงการพัฒนาสังคมแห่งความเท่าเทียมด้วย ICT กระทรวงเทคโนโลยีสารสนเทศและการสื่อสาร," วิทยาศาสตร์มหาบัณฑิต สาขาวิชาเทคโนโลยีคอมพิวเตอร์และการสื่อสาร ภาควิชาเทคโนโลยีคอมพิวเตอร์และการสื่อสาร บัณฑิตวิทยาลัย มหาวิทยาลัยธุรกิจบัณฑิต, 2553.
- [14] ปรีดี ปลื้มสำราญกิจ, "การตรวจสอบความสามารถในการเข้าถึงได้ทางเว็บ," *วารสารสารสนเทศศาสตร์* ปีที่ 29 ฉบับที่ 2 พ.ค.-ส.ค. 2554 หน้า 40-52.
- [15] ส่วนงานดัชนีและสำรวจสำนักนโยบายและส่งเสริมธุรกรรมทางอิเล็กทรอนิกส์ และสำนักงานพัฒนาธุรกรรมทางอิเล็กทรอนิกส์, "รายงานผลการสำรวจพฤติกรรมผู้ใช้อินเทอร์เน็ตในประเทศไทย ปี 2556," (ออนไลน์) สืบค้นวันที่ 11 มีนาคม 2557 จาก : http://www.etda.or.th/InternetUserprofile2013/TH_InternetUserProfile2013.pdf.

การประยุกต์ใช้การวัดระยะการแก้ไขกับข้อความภาษาไทย

Implementation of Edit Distance with Thai Language

เมธาวี จินดารักษ์ (Metawee Jindarak)¹ และสุภกิจ นุตยะสกุล (Supakit Nootyaskool)²

สาขาวิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

¹metawee.j@gmail.com, ²supakit@it.kmitl.ac.th

บทคัดย่อ

งานวิจัยนี้มีจุดประสงค์เพื่อศึกษาวิธีการในการวัดระยะการแก้ไข (Edit Distance) และพิสูจน์หาวิธีการที่เหมาะสมในการนำวิธีดังกล่าวมาประยุกต์ใช้กับภาษาไทย โดยได้ใช้วิธีการวัดระยะการแก้ไขมาประยุกต์ใช้ร่วมกับวิธีการตัดคำสำหรับภาษาไทย โดยงานวิจัยนี้ได้้นำ Levenshtein Distance ซึ่งเป็นวิธี Edit Distance ที่เป็นที่นิยม ไปทำการทดสอบเปรียบเทียบกับวิธี Needleman-Wunsch Algorithm และ Hamming Distance เพื่อเปรียบเทียบว่าวิธีการดังกล่าวเหมาะสำหรับการประยุกต์ใช้กับภาษาไทยหรือไม่ ซึ่งผลที่ได้จากการทดสอบนั้น ได้แสดงให้เห็นว่า Levenshtein Distance เป็นวิธีที่มีความเหมาะสมที่จะนำมาประยุกต์ใช้กับภาษาไทยเมื่อเทียบกับวิธีการอื่นๆ ที่นำมาเปรียบเทียบกัน เนื่องจาก Levenshtein Distance จะบอกได้ถึงจำนวนครั้งของการเปลี่ยนแปลงสำหรับการแก้ไขข้อความ ซึ่งต่างจาก Needleman-Wunsch Algorithm ที่บอกได้แค่ความเหมือนหรือต่างกันเพียงใด และ Levenshtein Distance ยังมีข้อจำกัดในการใช้งานน้อยกว่า Hamming Distance ด้วย

คำสำคัญ: การวัดระยะการแก้ไข การเปรียบเทียบข้อความภาษาไทย ระยะทางเลเวนชเต็น

Abstract

This research objective is to prove about how to implement edit distance to suit for Thai language. The research experiment is using Thai word segmentation method combine with various edit distance method such as Levenshtein Distance, Needleman-Wunsch Algorithm and Hamming Distance to find out what is the most

suitable edit distance method for using with Thai language. The result of the experiment shows that Levenshtein Distance is most suitable for implement with Thai language when comparing with other two edit distance methods because Levenshtein Distance can show the value of difference between two sentences that Needleman-Wunsch Algorithm and Hamming Distance cannot show that value.

Keywords: Edit Distance, Thai Sentence Comparison, Levenshtein Distance

1. บทนำ

ในการเปรียบเทียบความแตกต่างของข้อความ และรูปภาพ หรือรูปทรงต่างๆ ในทางคอมพิวเตอร์นั้น มีวิธีการที่นำมาใช้ในการเปรียบเทียบได้หลายวิธี ซึ่งวิธีการหนึ่งที่ถูกนำมาใช้คือ การวัดระยะการแก้ไข โดยการวัดระยะการแก้ไขนั้น ก็มีผู้คิดค้นอัลกอริทึมที่แตกต่างกันออกไปหลายวิธี โดยแต่ละวิธีจะมีวิธีการในการวัดระยะ และข้อจำกัดในการวัดระยะที่แตกต่างกัน เช่น การวัดระยะด้วยวิธี Hamming Distance จะเป็นการวัดระยะการแก้ไขข้อความที่มีจำนวนตัวอักษรเท่ากัน และเปรียบเทียบเฉพาะตัวอักษรที่ตำแหน่งที่ตรงกันเท่านั้น ซึ่งวิธีนี้ใช้ในการทำ Hamming Code สำหรับแก้ไขข้อมูลที่ผิดพลาด (Error Correction) จากการส่งข้อมูลต้นทางไปยังปลายทางในระบบเครือข่าย[1] เป็นต้น แต่ถ้าจะกล่าวถึงวิธีการวัดระยะการแก้ไขที่เป็นที่นิยม และรู้จักแพร่หลายคือ Levenshtein Distance ซึ่งเป็นวิธีการที่ใช้ในการวัดจำนวนครั้งของการแก้ไขข้อความ จากคำหนึ่งไปเป็นอีกคำหนึ่ง ด้วยการนับจำนวนครั้งของการ

แทรก การตัดออก และการแทนที่ตัวอักษร ในข้อความนั้นๆ ว่า ใช้จำนวนครั้งเท่าใดในการแก้ไขข้อความ

วิธีการวัดระยะการแก้ไขนั้นมักถูกนำไปใช้ในการเปรียบเทียบหาความแตกต่างระหว่างสิ่งของสองสิ่ง ซึ่งได้มีผู้นำทฤษฎีดังกล่าวไปทำการวิจัยในด้านต่างๆ มากมาย เช่น การนำไปประยุกต์ใช้ในการจับคู่รูปทรงที่มีลักษณะเหมือนกัน หรือใกล้เคียงกันโดยอาศัยหลักการ Shock Graph ที่นำเสนอโดย Philip N. Klein และคณะ [2] หรือการจำแนกหมวดหมู่ไวรัสคอมพิวเตอร์โดยอัตโนมัติ ด้วยการนำโครงสร้างหลักของชุดคำสั่งของไวรัสมาจำแนกด้วยวิธีการวัดระยะการแก้ไข ซึ่งคิดค้นขึ้นโดย Marius Gheorghescu จากบริษัทไมโครซอฟท์ [3] หรือการเปรียบเทียบความแตกต่างของวิดีโอคลิปด้วยการใช้ MPEG-7 Descriptor และวิธีการวัดระยะการแก้ไขที่นำเสนอโดย Marco Bertini และคณะ จากมหาวิทยาลัยฟลอเรนซ์ [4] เป็นต้น

2. การวัดระยะการแก้ไขข้อความ (Edit Distance)

การวัดระยะการแก้ไข (Edit Distance) เป็นวิธีการในการวัดหาความต่างกันระหว่างชุดตัวอักษรสองชุด ระหว่างชุดแรกที่เป็นชุดต้นแบบ และชุดที่สองที่เป็นชุดเปรียบเทียบ โดยค่าความแตกต่างจะวัดจากจำนวนครั้งของการที่จะต้องทำการแทรก ตัดออก และแทนที่ตัวอักษรในชุดตัวอักษรที่นำมาเปรียบเทียบจนกระทั่งชุดตัวอักษรที่นำมาเปรียบเทียบ (ชุดที่สอง) มีลักษณะเหมือนกับตัวอักษรชุดต้นแบบ (ชุดแรก) ทุกประการ ซึ่งโดยทั่วไปการวัดระยะการแก้ไขจะอ้างอิงถึงระยะทางเลเวนชไตน์ (Levenshtein Distance) ซึ่งเป็นทฤษฎีที่เกี่ยวข้องกับการวัดระยะการแก้ไขที่ถูกตั้งชื่อตาม Vladimir Levenshtein ผู้ที่ปรับปรุงชุดอัลกอริทึมชุดนี้ [5]

2.1 วิธีการของ Levenshtein Distance

ขั้นตอนวิธีการ Levenshtein Distance จะเป็นการนำชุดตัวอักษร 2 ชุด มาเปรียบเทียบจำนวนความแตกต่างกัน โดยจะพิจารณาจาก 3 รูปแบบ คือ

การแทรก : เป็นการนำเอาตัวอักษรตัวใดๆ มาเพื่อให้ชุดตัวอักษรชุดนั้นเหมือนกับอีกชุดตัวอักษรหนึ่งในภายหลัง เช่น MAN เป็น MAIN จะแทรกตัว I ให้กับ MAN เพื่อให้ MAN กลายเป็น MAIN เป็นต้น

การตัดออก : เป็นการตัดตัวอักษรออกครั้งละ 1 ตัว จากชุดตัวอักษรชุดหนึ่ง เพื่อให้ชุดตัวอักษรชุดนั้นเหมือนกับอีกชุดตัวอักษรหนึ่งในภายหลัง เช่น GOD เป็น GO จะตัดตัว D ออกเพื่อให้ GOD กลายเป็น GO เป็นต้น

การแทนที่ : เป็นการนำตัวอักษรของชุดตัวอักษรหนึ่งไปแทนตัวอักษรของอีกชุดตัวอักษรหนึ่ง เพื่อให้ชุดตัวอักษรชุดนั้นเหมือนกับอีกชุดตัวอักษรหนึ่งในภายหลัง เช่น CHAT เป็น THAT จะแทนที่ C ด้วย T เพื่อให้ CHAT กลายเป็น THAT หรือ อาจมองในทางกลับกันก็ได้ เป็นต้น [6]

2.2 ตัวอย่างของ Levenshtein Distance

ตารางที่ 1: ตารางการคำนวณหาค่าระยะการแก้ไขระหว่างคำว่า ENGLISH และ THAI

	#	T	H	A	I
#	0	1	2	3	4
E	1	1	2	3	4
N	2	2	2	3	4
G	3	3	3	3	4
L	4	4	4	4	4
I	5	5	5	5	4
S	6	6	6	6	5
H	7	7	6	7	6

จากตัวอย่างจะเห็นว่า คำว่า ENGLISH และคำว่า THAI มีค่าความต่างกันอยู่คือ 6 โดยได้มาจาก

- การตัดอักษร G, S, H ของคำว่า ENGLISH ออก โดยดำเนินการ 3 ครั้ง

- การแทนที่อักษร E ด้วยอักษร T, อักษร N ด้วยอักษร H และอักษร L ด้วยอักษร A โดยดำเนินการ 3 ครั้ง

ดังนั้นจะเห็นว่า จากการแก้ไขคำว่า ENGLISH ให้เป็นคำว่า THAI จะต้องดำเนินการตัดอักษร 3 ครั้ง และแทนที่ตัวอักษรด้วยตัวอักษรอื่นอีก 3 ครั้ง รวมเป็นการดำเนินการทั้งหมด 6 ครั้ง จึงหมายความว่า ทั้งสองคำมีค่าความแตกต่างกันอยู่เท่ากับ 6

2.3 รหัสเทียมของ Levenshtein Distance

```

int LevenshteinDistance(char s[1..m], char t[1..n])
declare int d[0..m, 0..n]

for i from 0 to m
  d[i, 0] := i
for j from 0 to n
  d[0, j] := j

for j from 1 to n
  for i from 1 to m
    if s[i] = t[j] then
      d[i, j] := d[i-1, j-1]
    else
      d[i, j] := minimum
      (
        d[i-1, j] + 1 // การตัดออก
        d[i, j-1] + 1 // การแทรก
        d[i-1, j-1] + 1 // การแทนที่
      )
  )
return d[m, n]
    
```

ภาพที่ 1: รหัสเทียมของ Levenshtien Distance

3. การประยุกต์ใช้การวัดระยะการแก้ไขกับข้อความภาษาไทย

ภาษาไทยเป็นภาษามาตรฐานภาษาหนึ่งที่ใช้งานบนคอมพิวเตอร์ โดยภาษาไทยนั้นมีตัวอักษรที่ใช้งานบนคอมพิวเตอร์ทั้งสิ้น 87 ตัว และมีรหัสสำหรับอักขระไทยที่ใช้บนคอมพิวเตอร์ คือ TIS-620 ซึ่งเป็นรหัสสำหรับอักขระไทยในตารางรหัส EBCDIC (Extended Binary Coded Decimal Interchange Code) ที่มีขนาด 8 บิต ซึ่งภายหลังยังได้มีการพัฒนาไปใช้ในรหัส Unicode ด้วย [7]

ตารางที่ 2: ตารางแสดงรหัสภาษาไทยสำหรับคอมพิวเตอร์ TIS-620

	x0	x1	x2	x3	x4	x5	x6	x7	x8	x9	xA	xB	xC	xD	xE	xF
0x	ไมโคร															
1x	ไมโคร															
2x	SP	!	"	#	\$	%	&	'	()	*	+	,	-	.	/
3x	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
4x	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
5x	P	Q	R	S	T	U	V	W	X	Y	Z	[\]	^	_
6x	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
7x	p	q	r	s	t	u	v	w	x	y	z	{		}	~	
8x	ไมโคร															
9x	ไมโคร															
Ax	ก	ข	ช	ค	ด	ต	ถ	ท	ธ	ด	ช	ฅ	ฌ	ญ	ฎ	ฏ
Bx	ร	ล	ฌ	ฌ	ด	ด	ถ	ท	ธ	น	บ	ป	ผ	ฝ	พ	ฟ
Cx	ภ	ม	ย	ร	ฤ	ล	ฎ	ว	ศ	ษ	ส	ห	ฬ	อ	ฮ	า
Dx	ะ	ั	ำ	อ	ิ	ี	ึ	ุ	ู	ุ	ุ	ุ	ุ	ุ	ุ	ุ
Ex	เ	แ	โ	ใ	ใ	า	า	อ	อ	อ	อ	อ	อ	อ	อ	อ
Fx	0	1	2	3	4	5	6	7	8	9	ฯ	๗				

3.1 ปัญหาของการใช้ Edit Distance กับภาษาไทย

เนื่องด้วยภาษาไทยนั้นจะมีลักษณะของข้อความที่มีตัวอักษรเขียนติดกันทั้งหมด และเมื่อจบประโยคจึงมีการเว้นวรรค ซึ่งไม่มีโครงสร้างของคำและประโยคที่ชัดเจนแบบ

ภาษาอังกฤษที่จะมีการเว้นวรรคเพื่อแบ่งคำ และมีการใส่จุด (.) เมื่อสิ้นสุดประโยค ทำให้ภาษาไทยนั้นไม่สามารถที่จะนำข้อความทั้งสองมาเปรียบเทียบกันได้ เพราะไม่มีการแบ่งขอบเขตของคำที่ชัดเจน เพราะหากไม่มีการแบ่งขอบเขตของคำในการเปรียบเทียบจะเกิดความคลาดเคลื่อน เนื่องจากระบบจะไม่สามารถจำแนกได้ว่าส่วนใดเป็นส่วนที่ผิด หรือส่วนใดถูกต้อง ทำให้ต้องมีการตัดแบ่งคำในข้อความออกมาเป็นชุดของคำก่อนที่จะมาทำการวัดระยะการแก้ไขของข้อความ ซึ่งวิธีในการตัดคำภาษาไทยนั้นก็สามารถทำได้หลายวิธี

3.2 รูปแบบของการตัดคำภาษาไทย

การตัดคำแบบยาวที่สุด (Longest Matching) : การตัดคำวิธีนี้ทำโดยเริ่มจากตัวอักษรซ้ายสุดของข้อความไปยังตัวอักษรถัดไป จนกว่าจะพบคำที่มีอยู่ในพจนานุกรม หลังจากนั้นก็ค้นหาคำถัดไปจนกว่าจะจบข้อความ หากพบว่าเป็นคำในพจนานุกรมจากจุดเริ่มต้นเดียวกัน จะเลือกเอาคำที่ยาวที่สุด เช่น การแบ่งคำในประโยค "คนตากลมนอนตากลม" จะเริ่มจากตัวอักษร ค แล้วค้นหาตัวอักษรถัดไปและนำมาเปรียบเทียบคำในพจนานุกรม โดยคำแรกที่แบ่งได้คือ "คน" และตัวอักษรถัดไปคือ ด จากตัวอักษรนี้จะได้คำว่า "ตา" กับคำว่า "ตาก" ซึ่งวิธีนี้ให้เลือกคำที่ยาวที่สุดที่ค้นพบจึงเลือกคำว่า "ตาก" จากนั้นก็จะค้นหาและเปรียบเทียบต่อไป ซึ่งผลลัพธ์ที่ออกมาคือ "คน ตาก ลม นอน ตาก ลม"

การตัดคำแบบสอดคล้องมากที่สุด (Maximal Matching) : เป็นการหาวิธีในการตัดคำที่เป็นไปได้ทั้งหมด เช่น ข้อความว่า "ไปหามเหสี" ซึ่งสามารถตัดคำได้ 2 แบบ คือ "ไป หาม เห สี" และ "ไป หาม เหสี" โดยที่วิธีการนี้จะเลือกเอาข้อความที่แบ่งแล้วได้จำนวนค่าน้อยที่สุด คือ "ไป หาม เหสี" ซึ่งมีจำนวนคำที่ตัดได้ 3 คำ ในขณะที่ "ไป หาม เห สี" มี 4 คำ

การตัดคำแบบคำนวณเชิงสถิติเพื่อหาความเป็นไปได้ (Probabilistic Model) : วิธีนี้นำเอาค่าสถิติการเกิดของคำและลำดับของหน้าที่ของคำ (Part of speech) มาช่วยในการคำนวณหาความน่าจะเป็น เพื่อใช้เลือกแบบที่มีโอกาสการเกิดมากที่สุด แต่วิธีนี้มีข้อจำกัด คือ ต้องมีฐานข้อมูลที่มีการตัดคำที่ถูกต้อง และต้องกำหนดหน้าที่ของคำ เพื่อที่จะได้นำไปใช้ในการสร้างสถิติ

การตัดคำแบบใช้คุณลักษณะ (Feature-Based Approach) : วิธีการนี้จะพิจารณาจากบริบทของคำ (Context words) และการเกิดร่วมกันหรือหน้าที่ของคำ (Collocation) เข้ามาช่วยในการตัดคำ เช่น "ตากลม" ถ้าพบคำว่า "โต" ในบริบทก็จะสามารถตัดคำได้ว่า "ตา กลม" ซึ่งวิธีนี้จำเป็นที่จะต้องมีความรู้ข้อมูลเป็นจำนวนมาก และจะต้องมีการเรียนรู้การสร้างคำในบริบทหรือการเกิดร่วมกันของคำแต่ละคำ เพื่อให้มีข้อมูลที่นำมาใช้ในการตัดคำ [8]

4. การทดสอบการวัดระยะการแก้ไขกับข้อความภาษาไทย

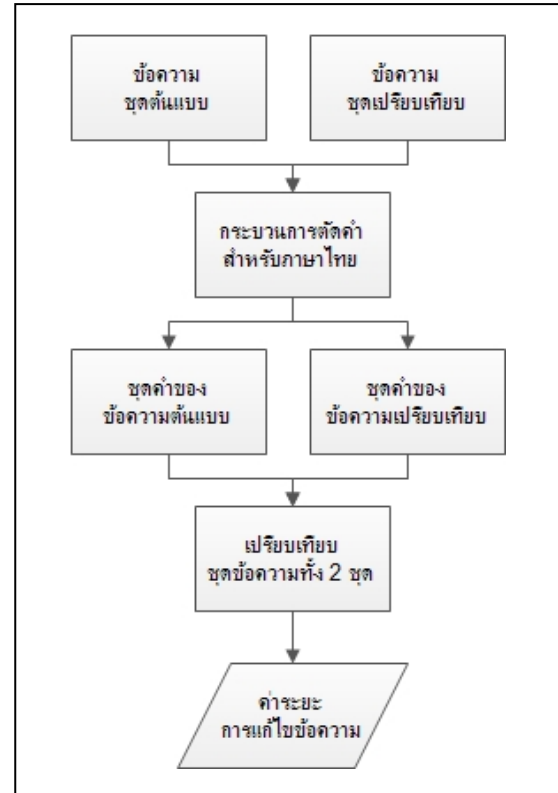
สำหรับการทดสอบการวัดระยะการแก้ไขข้อความด้วยวิธี Edit Distance กับข้อความภาษาไทยนั้น ได้มีการทดสอบโดยใช้กลุ่มของข้อความภาษาไทยตัวอย่างจำนวนหนึ่ง มาทำการทดสอบบนโปรแกรมที่ถูกพัฒนาขึ้น เพื่อคำนวณหาระยะในการแก้ไขข้อความสำหรับภาษาไทย เพื่อเปรียบเทียบหาวิธีการวัดระยะที่เหมาะสมสำหรับข้อความภาษาไทย ซึ่งได้วิธีต่างๆ มาใช้ในการเปรียบเทียบกัน 3 วิธี คือ Levenshtein Distance, Hamming Distance และ Needleman-Wunsch Algorithm [9] โดยที่โปรแกรมจะคำนวณหาค่าระยะการแก้ไขข้อความจำนวนคำที่แบ่งได้ รูปแบบของคำที่แบ่ง รวมไปถึงระยะเวลาที่ใช้ในการคำนวณ

การทำงานของโปรแกรมนั้นจะแบ่งขั้นตอนหลักๆภายในโปรแกรมได้ 2 ส่วน คือ ขั้นตอนในการตัดแบ่งข้อความออกเป็นคำ และขั้นตอนในการนำคำแต่ละคำของทั้งสองข้อความมาทำการเปรียบเทียบกันด้วยวิธี Edit Distance

ขั้นตอนในการตัดแบ่งข้อความออกเป็นคำนั้นได้ใช้โปรแกรม SWaTh (Smart Word Analysis for Thai) ซึ่งเป็นโปรแกรมสำหรับตัดคำภาษาไทยที่พัฒนาขึ้นโดย ไพศาล เจริญพรสวัสดิ์ [10] โดยที่โปรแกรมดังกล่าวจะใช้ลักษณะการตัดคำภาษาไทยหลายแบบผสมกัน ได้แก่ การตัดคำแบบยาวที่สุด (Longest Matching) และการตัดคำแบบสอดคล้องมากที่สุด (Maximal Matching) ร่วมกับการใช้ค่าสถิติการเกิดของคำและลำดับของหน้าที่ของคำ (Part of speech) เข้ามาช่วยในการตัดคำ

สำหรับขั้นตอนในการนำคำแต่ละคำของทั้งสองข้อความมาทำการเปรียบเทียบกันนั้น จะเป็นการนำคำแต่ละคำของ

ข้อความทั้งสองมาเปรียบเทียบกันแบบคำต่อคำ และเรียงตามลำดับ เพื่อหาค่าในการแก้ไขข้อความ โดยการนำคำที่หนึ่งของข้อความที่หนึ่ง มาเปรียบเทียบกับคำที่หนึ่งของข้อความที่สอง และทำเช่นนี้ต่อไปจนกระทั่งถึงคำสุดท้ายของข้อความ



ภาพที่ 2: กระบวนการในการวัดระยะแก้ไขข้อความภาษาไทย

4.1 วิธีการทดสอบการวัดระยะการแก้ไขข้อความภาษาไทย

ในการทดสอบการวัดระยะการแก้ไขข้อความภาษาไทยนั้น จะใช้วิธีการต่างกัน 3 วิธี มาใช้ในการวัดหาระยะการแก้ไข โดยแต่ละวิธีนั้นจะมีลักษณะและข้อจำกัดที่แตกต่างกันดังนี้

- Levenshtein Distance เป็นการวัดหาค่าความแตกต่างระหว่างข้อความสองชุด โดยนับจำนวนครั้งของการแทรก ตัดออก และแทนที่ตัวอักษร ซึ่งในการกระทำดังกล่าวแต่ละครั้งคะแนนในการวัดค่าจะเท่ากับ +1 ซึ่งผลคะแนนสุดท้ายที่วัดได้จะเป็นผลรวมของจำนวนครั้งที่ใช้ในการแก้ไขข้อความ

- Hamming Distance เป็นการวัดหาค่าความแตกต่างระหว่างข้อความสองชุด โดยที่ข้อความทั้งสองต้องมีจำนวนตัวอักษรที่เท่ากัน โดยการเปรียบเทียบจะเปรียบเทียบแบบตัวอักษรต่อตัวอักษร จากตัวอักษรตัวแรกจนถึงตัวอักษรตัวสุดท้าย หากพบว่าตัวอักษรต่างกันคะแนนในการวัดค่าจะเท่า

+1 ซึ่งผลคะแนนสุดท้ายที่วัดได้ จะเป็นผลรวมของจำนวนตัวอักษรที่มีความแตกต่างกันทั้งหมด

- Needleman-Wunsch Algorithm เป็นการวัดค่าความเหมือนระหว่างข้อความสองชุด โดยเป็นการนำคำของข้อความทั้งสองมาหาค่าที่ใช้ในการจัดเรียงตัวอักษร โดยที่เมื่อนำคำมาเปรียบเทียบกันแล้วตำแหน่งที่เปรียบเทียบกันเป็นเป็นตัวอักษรทั้งคู่และเป็นตัวอักษรตัวเดียวกัน ค่าที่ได้จะเป็น +2 แต่ถ้าเป็นตัวอักษรคนละตัวกัน ค่าที่ได้จะเป็น -1 และหากตำแหน่งที่เปรียบเทียบเป็นช่องว่างกับตัวอักษร ค่าที่ได้จะเป็น -2

4.2 ผลการทดสอบการวัดระยะการแก้ไขข้อความภาษาไทย

ในการทดสอบการวัดระยะการแก้ไขข้อความภาษาไทยได้ใช้ชุดข้อความภาษาไทยตัวอย่าง นำมาเปรียบเทียบกันเป็นคู่จำนวน 5 ชุดข้อความ ดังนี้

ชุดที่ 1: “การเรียน ” เปรียบเทียบกับ “ การเกษตร ”

ชุดที่ 2: “ ตกกระไดพลอยโจน ” เปรียบเทียบกับ “ ตีวัวกระทบลาด ”

ชุดที่ 3: “ รู้หลบเป็นปีก รู้หลิกเป็นหาง ” เปรียบเทียบกับ “ รักดีหามจ้าว รักชั่วหามเสา ”

ชุดที่ 4: “ ศิแแดงแห้วชอว์เบะท่าไปเชลชี ” เปรียบเทียบกับ “ ศิควัวชัยรับขวัญมาตาเปดรงทุบคาร์ดิฟี่ ”

ชุดที่ 5: “ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ” เปรียบเทียบกับ “ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้า พระนครเหนือ ”

โดยผลการทดสอบจากการเปรียบเทียบชุดข้อความตัวอย่างข้างต้น จะได้จากค่าจากการเปรียบเทียบตามตารางดังต่อไปนี้

ตารางที่ 3: ตารางแสดงผลการทดสอบการวัดระยะการแก้ไขข้อความภาษาไทยด้วยวิธีการต่างๆ

ชุดข้อความ	Hamming		Levenshtein			Needleman-Wunsch		
	Score A	Time	Score A	Score B	Time	Score A	Score B	Time
1	4	1	4	4	83	4	4	76
2	13	1	12	25	85	-5	-4	76
3	N/A	N/A	24	26	97	-18	-14	88
4	N/A	N/A	34	37	98	-22	-15	81
5	N/A	N/A	29	53	82	13	-30	84

Score A : ค่าที่ได้จากการวัดระยะการแก้ไข โดยไม่ตัดแบ่งคำ
 Score B : ค่าที่ได้จากการวัดระยะการแก้ไข โดยมีการตัดแบ่งคำ
 Time : ระยะเวลาที่ใช้ในการคำนวณหาในการวัดระยะการแก้ไข (Millisecond)

จากตารางแสดงผลการทดสอบการวัดระยะการแก้ไขข้อความภาษาไทยข้างต้น จะเห็นว่าในการวัดระยะการแก้ไขข้อความภาษาไทยนั้น วิธีการที่ใช้เวลาในการวัดระยะการวัดเร็วที่สุดจะเป็น Hamming Distance, Needleman-Wunsch

Algorithm และ Levenshtein Distance ตามลำดับ แต่ Hamming Distance นั้นจะไม่สามารถตัดแบ่งคำก่อนการวัดระยะได้ เพราะจะทำให้จำนวนตัวอักษรในการนำมาเปรียบเทียบไม่เท่ากัน จึงทำให้ Hamming Distance ต้องเปรียบเทียบทั้งข้อความ โดยที่ทั้งสองข้อความต้องมีจำนวนตัวอักษรเท่ากันด้วย

สำหรับ Levenshtein Distance และ Needleman-Wunsch Algorithm นั้นสามารถตัดแบ่งคำก่อนการวัดค่าได้ แต่หากจำนวนของคำที่นำมาเปรียบเทียบกันมีไม่เท่ากัน ค่าที่ได้จาก Needleman-Wunsch Algorithm จะมีความคลาดเคลื่อนเนื่องจากค่าที่วัดได้จากตัวอักษรกับคำที่ไม่มีข้อความ (‘’) จะได้ค่าเป็น 0 ซึ่งคลาดเคลื่อนจากความเป็นจริง แต่ Levenshtein Distance นั้นจะวัดค่าได้ถูกต้องและแม่นยำกว่า

5. สรุปผลการทดสอบการวัดระยะการแก้ไขข้อความภาษาไทย

ในการประยุกต์ใช้การวัดระยะการแก้ไขกับข้อความภาษาไทยนั้น Levenshtein Distance ถือเป็นวิธีที่เหมาะสมกับข้อความภาษาไทย เนื่องจากไม่มีปัญหาเรื่องค่าที่คลาดเคลื่อนจากการตัดแบ่งคำที่ได้จำนวนของคำไม่เท่ากัน และยังได้ค่าที่ตรงตัว โดยไม่ต้องไปแปลผลของค่าที่ได้ ซึ่งจะต่างจาก Needleman-Wunsch Algorithm ที่เป็นวิธีในการวัดค่าความเหมือนกันของข้อความ โดยค่าที่ได้จากวิธีนี้จะบอกได้แค่ข้อความทั้งสองที่นำมาเปรียบเทียบกันมีความเหมือนกันมากหรือน้อยเพียงใด ขึ้นอยู่กับค่าที่วัดได้ว่าเป็นบวกหรือลบ

สำหรับ Needleman-Wunsch Algorithm นี้จะเหมาะกับการวัดระยะการแก้ไขข้อความในลักษณะของคำมากกว่าจะมาวัดแบบทั้งประโยค เพราะหากมีการตัดแบ่งคำแล้ว ค่าที่ได้จาก Needleman-Wunsch Algorithm จะมีความคลาดเคลื่อน ทำให้ผลลัพธ์ที่ออกมาไม่ถูกต้อง

ซึ่งจากผลการทดสอบดังกล่าว สามารถแนะนำแนวทางของการทดสอบนี้ไปต่อยอดเพื่อประยุกต์ใช้กับภาษาไทยได้ เช่น การทำระบบการแก้ไขคำผิดสำหรับภาษาไทยโดยอัตโนมัติ (Auto Correction) เป็นต้น จะทำให้เกิดประโยชน์มากยิ่งขึ้นต่อการใช้ภาษาไทยบนระบบคอมพิวเตอร์

เอกสารอ้างอิง

- [1] Department of Electrical and Computer Engineering - University of New Brunswick. n.p. *Error Correction and the Hamming Code*. [Online] Available: <http://www.ee.unb.ca/tervo/ee4253/hamming.shtml>
- [2] Philip N. Klein, Thomas B. Sebastian and Benjamin B. Kimia. 2001. "Shape matching using edit-distance: an implementation." 781-790. In *SODA '01 Proceedings of the twelfth annual ACM-SIAM symposium on Discrete algorithms*, Philadelphia: SIAM.
- [3] Marius Gheorghescu. 2006. "An Automated Virus Classification System." In *Virus Bulletin Conference October 2005*, Washington: Virus Bulletin.
- [4] Marco Bertini, Alberto Del Bimbo and Walter Nunziati. 2006. "Video Clip Matching Using MPEG-7 Descriptors and Edit Distance." 133-142. In *5th International Conference, CIVR 2006*, New York: Springer Berlin Heidelberg
- [5] Levenshtein VI (1966). "Binary codes capable of correcting deletions, insertions, and reversals". *Soviet Physics Doklady* 10: 707–10.
- [6] Dan Jurafsky. n.p. *Minimum Edit Distance*. [Online] Available: www.stanford.edu/class/cs124/lec/med.pdf
- [7] สำนักงานมาตรฐานผลิตภัณฑ์อุตสาหกรรม. ม.ป.ป. *รหัสสำหรับอักขระไทยที่ใช้กับคอมพิวเตอร์*. [Online] Available: <http://www.nectec.or.th/it-standards/std620/std620.htm>
- [8] "โปรแกรมตัดคำภาษาไทย (Thai Word Segmentation)." *สารานุกรมไทยสำหรับเยาวชน โดยพระราชประสงค์ในพระบาทสมเด็จพระเจ้าอยู่หัว*. เล่มที่ 25
- [9] S. B. Needleman and C. D. Wunsch. 1969. "A General Method Applicable to the Search for Similarities in the Amino Acid Sequence of Two Proteins." *Department of Biochemistry, Northwestern University*.
- [10] Paisarn Charoenpornasawat. n.p. *SWATH - Thai Word Segmentation*. [Online] Available: <http://www.cs.cmu.edu/~paisarn/software.html>

การวิเคราะห์กาพย์ยานีสิบเอ็ดโดยใช้วิธีการตัดคำและฐานกฎ

A Eleven-Syllable Poetry Analysis using Word Segmentation and Rule-Based

อิสรานอล พลเกรีย์ (Isranol Polparee)¹ และ พุสดี บุญรอด (Pudsadee Boonrawd)²

สาขาวิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹isranol_boom@hotmail.com, ²pudsadee@kmutnb.ac.th

บทคัดย่อ

บทร้อยกรองมีหลายประเภทและมีฉันทลักษณ์ที่แตกต่างกัน ซึ่งกาพย์ยานีสิบเอ็ดเป็นประเภทหนึ่งของบทร้อยกรองที่วรรคหน้าและวรรคหลังมีจำนวนคำไม่เท่ากัน การสัมผัสมีทั้งระหว่างวรรคและระหว่างบท โดยการประพันธ์นั้นต้องเสียเวลาในการตรวจสอบความถูกต้องตามฉันทลักษณ์ ดังนั้นวัตถุประสงค์ของงานวิจัยนี้จึงได้ออกแบบขั้นตอนวิธีในการวิเคราะห์ฉันทลักษณ์ของกาพย์ยานีสิบเอ็ดโดยใช้วิธีการตัดคำ ฐานกฎ และพัฒนาโปรแกรมจำลองเพื่อใช้ในการทดลองการวัดความถูกต้อง (Accuracy) ของผลลัพธ์ที่ได้จากการวิเคราะห์เปรียบเทียบกับผลเฉลยมีค่าเท่ากับร้อยละ 83 การประเมินคุณภาพโปรแกรมจำลองจากผู้ใช้งานโดยภาพรวมอยู่ในระดับดีมาก จึงสรุปได้ว่าขั้นตอนวิธีที่พัฒนาขึ้นสามารถวิเคราะห์กาพย์ยานีสิบเอ็ดได้อย่างมีประสิทธิภาพ และโปรแกรมจำลองที่พัฒนาขึ้นสามารถนำไปประยุกต์ใช้งานได้จริง

คำสำคัญ: กาพย์ยานีสิบเอ็ด ฉันทลักษณ์ วิธีการตัดคำ ฐานกฎ

Abstract

There are many kinds of poetries with different prosodies. Eleven-syllable poetry called Gab Yani is kind of poetry with a number of words at the front and rear space. There is rhyme between the space and between each verse. When composing, it takes a lot of time to check for accuracy under the prosody rule. Therefore, this experimental system was designed to use word segmentation, rule-based techniques and the program simulation was developed to in

order to analyze the eleven-syllable poetry. The research divided the assessment into 2 aspects as follows, the accuracy of processes and comparison to methods with answers equivalent to 83 percent. The quality assessment of the program simulation was conducted by expert. In conclusion, the process, methods developed in this research used to analyze the eleven-syllable poetry successfully and simulation program developed can actually be applied to usable.

Keywords: Eleven-Syllable Poetry, Prosody, Word-Segmentation, Rule-Based

1. บทนำ

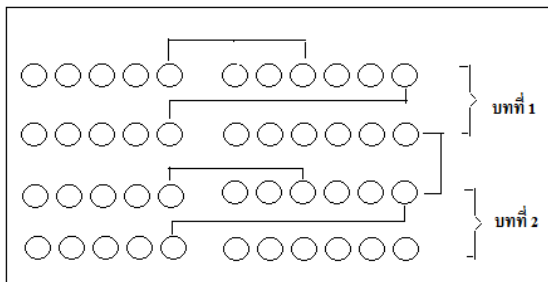
บทร้อยกรองเป็นการประพันธ์เรื่องราวต่าง ๆ ที่มีฉันทลักษณ์ทำให้เกิดความไพเราะ [1] เนื่องจากบทร้อยกรองมีหลายประเภทมีฉันทลักษณ์ที่เป็นรูปแบบเฉพาะ ซึ่งกาพย์ยานีสิบเอ็ดเป็นบทร้อยกรองประเภทหนึ่งที่วรรคหน้ามี 5 คำ วรรคหลังมี 6 คำ มีสัมผัสระหว่างวรรคและระหว่างบท ในการประพันธ์จำเป็นต้องได้รับการตรวจสอบความถูกต้องให้ตรงตามฉันทลักษณ์ การตรวจสอบนั้นอาจจะต้องใช้เวลานานหรือเกิดการผิดพลาดในการตรวจสอบได้ วัตถุประสงค์ของการทดลองเพื่อออกแบบการวิเคราะห์กาพย์ยานีสิบเอ็ดโดยใช้ฐานกฎและวิธีการตัดคำ เพื่อพัฒนาขั้นตอนวิธีและให้ได้มาซึ่งกรอบแนวคิดในการวิเคราะห์ที่ใช้ในการพัฒนาโปรแกรมจำลองเพื่อแก้ปัญหาและเป็นส่วนช่วยในการตรวจสอบฉันทลักษณ์ โดยในส่วนแรกอธิบายถึงปัญหาของงานวิจัย ส่วนที่สองแสดงทฤษฎีที่ใช้ในการแก้ปัญหา ส่วนที่สามอธิบายขั้นตอนวิธีและกรอบแนวคิดในการวิเคราะห์ฉันทลักษณ์

ของกาพย์ยานีสิบเอ็ด ส่วนที่สี่แสดงผลการวิจัย และส่วนสุดท้ายสรุปผลที่ได้จากการวิจัย

2. ทฤษฎีที่เกี่ยวข้อง

2.1 กาพย์ยานีสิบเอ็ด (Eleven-Syllable Poetry)

กาพย์ยานีสิบเอ็ดเป็นบทร้อยกรองประเภทหนึ่งในงานวิจัยนี้ได้นำเอามาเป็นกรณีศึกษาเพื่อใช้ในการทดลองวิเคราะห์และตรวจสอบ จึงได้นำเอาฉันทลักษณ์ของกาพย์ยานีสิบเอ็ดมาใช้เป็นฐานกฎในการวิเคราะห์ ฉันทลักษณ์ของกาพย์ยานีสิบเอ็ดแสดงดังรูปที่ 1



ภาพที่ 1: ฉันทลักษณ์ของกาพย์ยานีสิบเอ็ด

กาพย์ยานีสิบเอ็ดเป็นคำประพันธ์ประเภทกาพย์ คือ บทหนึ่งมี 2 บาท บาทหนึ่งมีสองวรรค วรรคแรกมี 5 คำวรรคหลังมี 6 คำ รวมเป็น 11 คำ จึงเรียกว่ากาพย์ยานีสิบเอ็ด สัมผัสของกาพย์ยานีสิบเอ็ด คือ คำสุดท้ายของวรรคที่หนึ่งสัมผัสกับคำที่ 3 ของวรรคที่สอง คำสุดท้ายของวรรคที่สองสัมผัสกับคำสุดท้ายของวรรคที่สาม และมีสัมผัสระหว่างบท คือ คำสุดท้ายของวรรคที่สี่สัมผัสกับคำสุดท้ายของวรรคที่สองของบทต่อไป การสัมผัสนี้จะทำให้กาพย์ยานีสิบเอ็ดมีความไพเราะมากยิ่งขึ้น [2] ในการสัมผัสมีสัมผัสบังคับ คือ สัมผัสนอกที่บทร้อยกรองทุกประเภทต้องมีและกำหนดให้ใช้แต่สัมผัสระหว่างบทนั้น

2.2 การตัดคำ (Word-Segmentation)

การตัดคำโดยใช้พจนานุกรม คือ พื้นฐานของระบบงานที่เกี่ยวข้องกับการประมวลผลภาษาธรรมชาติการตัดคำภาษาไทยนั้นจะทำได้ยากเนื่องจากไม่ทราบขอบเขตของประโยคเนื่องจากภาษาไทยไม่มีการเว้นวรรค มีคำก่าวมหรือมีคำที่ไม่อยู่ในพจนานุกรมทำให้ผลลัพธ์ของการตัดคำผิดพลาดส่งผลกระทบต่อผลลัพธ์ของระบบงาน [3] การตัดคำมีหลายวิธีที่ได้มี

การนำเสนอและมีการวิจัย เพื่อให้สามารถตัดคำในภาษาไทยให้ถูกต้องมากที่สุด จึงได้มีการนำเสนอวิธีการใช้พจนานุกรมมาช่วยในการตัดคำ แต่วิธีนี้จะมีข้อเสีย คือ จะตัดคำไม่ได้ถ้าคำนั้นไม่อยู่ในพจนานุกรม การตัดคำโดยใช้พจนานุกรม เป็นพื้นฐานของการตัดคำเพราะส่วนใหญ่คำจะอยู่ในพจนานุกรม การตัดคำภาษาไทยด้วยวิธี Maximal Matching เป็นหนึ่งในวิธีที่นิยมใช้กันซึ่งจะต้องมีพจนานุกรมที่มีคำศัพท์มากพอสมควร และถ้ามีคำที่ไม่อยู่ในพจนานุกรมจะทำให้การตัดคำผิดพลาด [4] และการตัดคำด้วยพจนานุกรมสามารถตัดคำได้ดีกว่าการตัดคำด้วยการเรียนรู้ของเครื่องบางวิธี เช่น Naïve Bayes (NB), Decision Tree, Support Vector Machine (SVM) [5] จึงนำเอาวิธีการตัดคำโดยใช้พจนานุกรมเข้ามาช่วยในการทำทดลอง

2.3 ฐานกฎ (Rule-Based)

ฐานกฎจะใช้การพัฒนากฎเกณฑ์หรือการกระทำที่มีเหตุมีผลให้อยู่ในรูปแบบของ IF-THEN ที่สามารถใช้สร้างแบบจำลองเพื่อวิเคราะห์และตัดสินใจได้ [6] ซึ่งฐานกฎจะมีความสัมพันธ์ของระบบกฎความรู้ในสภาพแวดล้อมขององค์ความรู้ [7] ฐานกฎเป็นรูปแบบขององค์ความรู้โดยปกติแล้วจะสร้างขึ้นจากองค์ความรู้ของมนุษย์ [8] ฐานกฎจึงเป็นส่วนที่มีความจำเป็นในการพัฒนาระบบที่เกี่ยวกับการวิเคราะห์การจัดหมวดหมู่ เป็นต้น

วิธีการแทนองค์ความรู้ด้วยกฎ กฎขององค์ความรู้โดยทั่วไปประกอบไปด้วย 3 ส่วน คือ ECA (Event-Condition-Action) [9] ซึ่ง E คือ ON event, C คือ IF condition, A คือ THEN action ในส่วนของ ON เป็นเหตุการณ์ IF เป็นเงื่อนไข และ THEN เป็นผลลัพธ์หรือการกระทำ การทำงานของ Rule-based system เริ่มต้นด้วยการสร้างกฎซึ่งบรรจุความรู้ที่เหมาะสมทั้งหมดและแปลงความรู้เหล่านั้นให้เป็นกฎที่อยู่ในรูปของคำสั่ง IF-THEN กฎขององค์ความรู้ไม่สามารถที่จะมีเงื่อนไขเดียวตลอดได้ อาจจะมีหลายเงื่อนไขให้เลือกกระทำ โดยการกระทำหลาย ๆ การกระทำสามารถให้ตัวเชื่อมเข้ามาช่วยได้ คือ AND หรือ OR [10] และผลลัพธ์ของการกระทำไม่จำเป็นที่จะต้องมียุทธวิธีเดียวอาจจะมีหลายผลลัพธ์ในเงื่อนไขเดียวได้เช่นกัน [11]

2.4 งานวิจัยที่เกี่ยวข้อง

ในการทดลองวิเคราะห์ภาพย้านิสิบเอ็ด ทฤษฎีที่นำมาประยุกต์ใช้มีอยู่ 2 ทฤษฎี คือ ฐานกฎ และวิธีการตัดคำ จึงได้ศึกษางานวิจัยที่มีความเกี่ยวข้องเพื่อความเข้าใจและสามารถนำมาประยุกต์ใช้กับงานได้อย่างถูกต้องและเหมาะสม

งานวิจัยของพัชรินทร์ [12] ได้ทำการเปรียบเทียบวิธีการตัดคำโดยใช้พจนานุกรมและการตัดคำโดยใช้การเรียนรู้ของเครื่องด้วยวิธี Conditional Random Fields (CRFs) จากงานวิจัยได้สรุปว่าการใช้การตัดคำโดยใช้การเรียนรู้ของเครื่องด้วยวิธีดังกล่าวมีประสิทธิภาพมากกว่าการตัดคำโดยใช้พจนานุกรม และยังกล่าวอีกว่าพจนานุกรมของ LEXiTRON มีประสิทธิภาพมากกว่าพจนานุกรมของ LEXEME ซึ่งสอดคล้องกับงานวิจัยของ Choochart [5] ที่ทำการวิจัยเปรียบเทียบวิธีการตัดคำโดยใช้การเรียนรู้ของเครื่อง 4 วิธี วิธีการตัดคำโดยใช้พจนานุกรม 2 วิธี ผลที่ได้คือ วิธีการตัดคำแบบ Conditional Random Field (CRFs) ให้ผลดีที่สุดสำหรับงานวิจัยนี้ และทำให้รู้ว่าการตัดคำโดยใช้พจนานุกรมถือเป็นวิธีการที่นิยมใช้ในการตัดคำภาษาไทย และพจนานุกรมของ LEXiTRON เป็นพจนานุกรมที่มีประสิทธิภาพที่ใช้ในการตัดคำ

3. วิธีการดำเนินการวิจัย

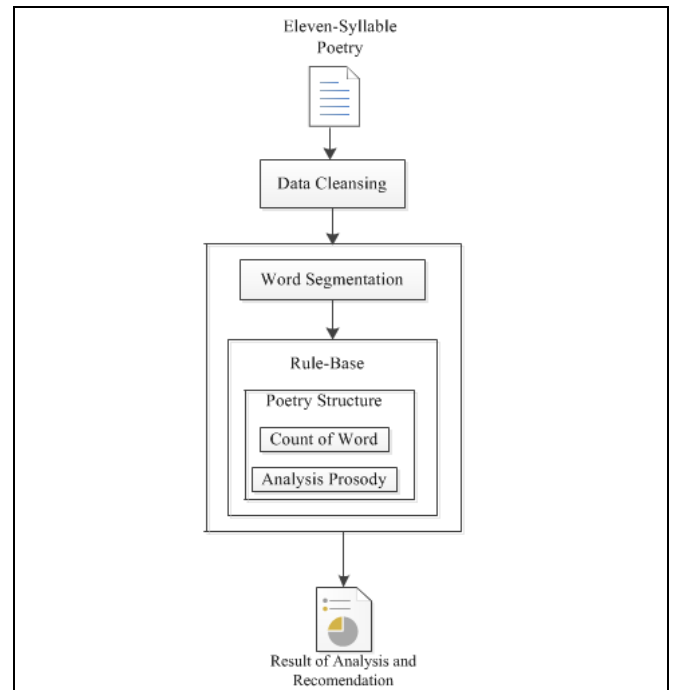
ในขั้นตอนการดำเนินการวิจัยนี้ได้ทำการรวบรวมและศึกษาปัญหาของการตรวจสอบทรัพยากรที่มีอยู่ จึงได้นำเอาฐานกฎและวิธีการตัดคำเข้ามาช่วยในการตรวจสอบและวิเคราะห์ทำให้สามารถพัฒนาขั้นตอนวิธีและได้มาซึ่งกรอบแนวคิดที่ใช้ในการพัฒนาโปรแกรมจำลองขึ้นมาใช้ในการทดลอง

จากที่ได้พัฒนาโปรแกรมจำลองโดยใช้วิธีการตัดคำและฐานกฎเพื่อใช้ในการวิเคราะห์ได้ทำการทดลองโดยการนำเอาภาพย้านิสิบเอ็ดมาเข้าสู่ขั้นตอนวิธีที่ได้พัฒนาขึ้นและทำการเก็บผลที่ได้เพื่อสรุปผลการวิจัยและเพื่อแก้ไขข้อบกพร่องของการวิเคราะห์ให้มีความถูกต้องมากยิ่งขึ้น เมื่อโปรแกรมจำลองสามารถวิเคราะห์ภาพย้านิสิบเอ็ดได้แล้วจึงให้ผู้ใช้งานได้ทดลองใช้และประเมินผล จากการพัฒนาโปรแกรมจำลอง

ขั้นตอนวิธีและกรอบแนวคิดสามารถแสดงได้ในหัวข้อ 3.1 และวิธีการประเมินผลแสดงในหัวข้อ 3.2

3.1 ขั้นตอนวิธีและกรอบแนวคิดของการวิเคราะห์

ขั้นตอนวิธีและกรอบแนวคิดถือเป็นส่วนสำคัญของการพัฒนาโปรแกรมจำลองที่ใช้ในการวิเคราะห์ ขั้นตอนที่สำคัญที่สุด คือ การตัดคำ และฐานกฎ ขั้นตอนวิธีและกรอบแนวคิดของการวิเคราะห์แสดงดังภาพที่ 2



ภาพที่ 2: ขั้นตอนวิธีและกรอบแนวคิดของการวิเคราะห์

ภาพที่ 2 แสดงขั้นตอนวิธีและกรอบแนวคิดของการวิเคราะห์ภาพย้านิสิบเอ็ดด้วยฐานกฎและวิธีการตัดคำ ซึ่งในการวิเคราะห์จะเริ่มจากการกรอกภาพย้านิสิบเอ็ดเข้าสู่กระบวนการวิเคราะห์ จากนั้นจะทำการเตรียมข้อมูลและส่งไปตัดคำด้วยวิธีการตัดคำที่พัฒนาขึ้น หลังจากตัดคำแล้วจะส่งผลการตัดคำที่ได้ไปยังกระบวนการวิเคราะห์ด้วยฐานกฎ และแสดงผลการวิเคราะห์ ในการแสดงผลนั้นจะแสดงผลของการตัดคำและผลการวิเคราะห์นั้นหลักเกณฑ์ด้วยฐานกฎ

ขั้นตอนการวิเคราะห์มีส่วนสำคัญ 2 ส่วน คือ

1. วิธีการตัดคำ

วิธีการตัดคำที่ใช้ได้นำเอาพจนานุกรมของ LEXiTRON มาใช้และแปลงคำศัพท์ให้เป็นคำศัพท์ที่มีพยางค์เดียว ในการตัดคำใช้วิธีการตัดคำแบบยาวที่สุด (Longest Matching) ในขั้นตอนการตัดคำจะรับภาพย้านิสิบเอ็ดทุกวรรคเข้ามาตัดคำ จากนั้นจะ

นำเอาผลการตัดคำที่ได้ไปเข้ากระบวนการวิเคราะห์ด้วยฐานกฎ ผลการตัดคำที่ได้จากวิธีการตัดคำแสดงตัวอย่างดังภาพที่ 3

คลื่น	คลั่ง	ทะ	เล	โถม	กระ	หน้า	โหม	ทะ	เล	คน
คลื่น	ซัด	ระ	บัด	ชล	เฉก	จะ	ล้าง	ๆ	อย่าง	ไร
อ้อม	แอ้ม	แค้น	เกลื่อน	อยู่	มิ	นึก	รู้	เหตุ	การณ์	ใด
ธ	ณี	พิ	โรธ	โย	จึง	เช่น	ฆ่า	ประ	ชา	กร

ภาพที่ 3: ผลการตัดคำ

2. ฐานกฎ

ฐานกฎเป็นขั้นตอนการวิเคราะห์ลักษณะของภาพยานีสืบเสาะที่ได้จากการตัดคำ จึงได้นำเอาฉันทลักษณ์มาเป็นฐานกฎของการวิเคราะห์ ฐานกฎจะแบ่งเป็น 2 ส่วน คือ

1. การนับจำนวนคำโดยการแบ่งการนับเป็นวรรคหน้า และวรรคหลัง ที่มีจำนวนคำไม่เท่ากัน คือ วรรคหน้า 5 คำ วรรคหลัง 6 คำ
2. การวิเคราะห์จุดสัมผัสที่มีทั้งระหว่างวรรคและสัมผัสระหว่างบทจากฉันทลักษณ์ของภาพยานีสืบเสาะสามารถแสดงเป็นฐานกฎได้ดังตารางที่ 1 ซึ่งจะเห็นว่าได้นำเอาตัวเชื่อม AND มาเป็นตัวเชื่อมระหว่างกฎย่อยในฐานกฎของการวิเคราะห์ เนื่องจากกฎทุกข้อต้องเป็นจริงทั้งหมด

ตารางที่ 1: ฐานกฎของภาพยานีสืบเสาะ

IF	<วรรคหน้า 5 คำ > AND <วรรคหลัง 6 คำ>
	AND <คำสุดท้ายของวรรคที่ 1 สัมผัสกับคำที่ 3 ของวรรคที่ 2>
	AND <คำสุดท้ายของวรรคที่ 2 สัมผัสกับคำสุดท้ายของวรรคที่ 3>
	AND <คำสุดท้ายของวรรคที่ 4 สัมผัสกับคำสุดท้ายของวรรคที่ 6>
	AND <คำสุดท้ายของวรรคที่ 5 สัมผัสกับคำที่ 3 ของวรรคที่ 6>
	AND <คำสุดท้ายของวรรคที่ 6 สัมผัสกับคำสุดท้ายของวรรคที่ 7>
THEN	<TRUE>

จากขั้นตอนในการวิเคราะห์ภาพยานีสืบเสาะทำให้ได้กรอบแนวคิดในการวิเคราะห์เพื่อนำไปพัฒนาโปรแกรมจำลองที่ใช้ในการทดลองการวิเคราะห์ในขั้นตอนของการตัดคำ และการวิเคราะห์ด้วยฐานกฎ ถือเป็นส่วนสำคัญ ถ้าวิธีการตัดคำที่

พัฒนาขึ้นตัดคำผิดพลาด หรือการออกแบบฐานกฎผิดจะทำให้การวิเคราะห์นั้นผิดไปด้วย

3.2 การประเมินผลโปรแกรมจำลองและขั้นตอนวิธี

การประเมินผลโปรแกรมจำลองที่ได้พัฒนาขึ้น ได้แบ่งการทดสอบออกเป็น 2 ประเภท คือ การประเมินผลตามฉันทลักษณ์ และ Functional Test

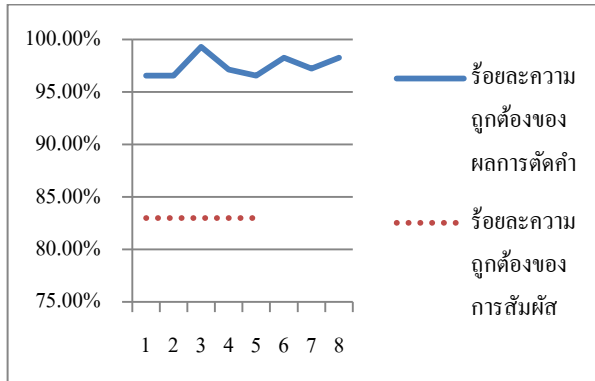
1. การประเมินผลตามฉันทลักษณ์ เป็นการประเมินผลขั้นตอนวิธีของการวิเคราะห์ที่ได้พัฒนาขึ้น ซึ่งจะครอบคลุมทั้งขั้นตอนวิธีในการวิเคราะห์ของฐานกฎและวิธีการตัดคำที่ได้พัฒนาขึ้น

2. Functional Test จะเป็นการประเมินผลโดยผู้ใช้ต่อโปรแกรมจำลองที่ได้พัฒนา ซึ่งความพึงพอใจนี้จะแสดงให้เห็นว่าการวิเคราะห์ภาพยานีสืบเสาะโดยใช้วิธีการตัดคำและฐานกฎสามารถช่วยวิเคราะห์ได้ตรงตามความต้องการของผู้ใช้

4. ผลการวิจัยและการอภิปรายผล

จากการทดลองวิเคราะห์ภาพยานีสืบเสาะด้วยฐานกฎได้มีการประเมินผลโดยการตรวจสอบความถูกต้องในการวิเคราะห์ด้วยฐานกฎ ความถูกต้องในการตัดคำที่ได้พัฒนาขึ้น และการตรวจสอบการสัมผัสตามฉันทลักษณ์ ซึ่งในการทดลองได้นำเอาภาพยานีสืบเสาะมาใช้ทดลองวิเคราะห์โดยโปรแกรมจำลอง

การประเมินผลการวิเคราะห์ได้นำภาพยานีสืบเสาะมาใช้ในการทดลองจำนวน 30 ตัวอย่าง โดยนำเอาภาพยานีสืบเสาะที่มีฉันทลักษณ์ถูกต้องมาทดลองกับโปรแกรมจำลอง ผลที่ได้จากการทดลองวิเคราะห์ภาพยานีสืบเสาะด้วยโปรแกรมจำลองโดยใช้วิธีการตัดคำที่พัฒนาขึ้นร่วมกับฐานกฎสามารถวิเคราะห์ถูกต้องประมาณร้อยละ 83 ในการทดลองวิเคราะห์ยังได้นำเสนอผลการตัดคำและผลการวิเคราะห์จุดสัมผัสตามฉันทลักษณ์ ซึ่งผลการทดลองแสดงดังภาพที่ 4

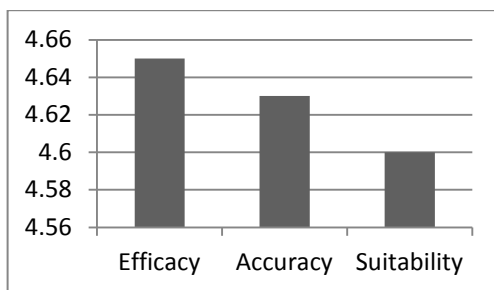


ภาพที่ 4: ร้อยละความถูกต้องของการตัดคำและการวิเคราะห์การสัมผัส

เส้น ไข่ปลา แสดงผลการตรวจสอบการสัมผัสของ กภย๋านี สิบเอ็ด ซึ่งสัมผัสของกภย๋านี สิบเอ็ดจะมีอยู่ 5 จุด มีร้อยละความถูกต้องของการวิเคราะห์เท่ากับ 83

เส้นทึบ แสดงผลการตัดคำที่ได้จากการตัดคำด้วยวิธีการตัดคำที่ ยาวที่สุด ซึ่งในกภย๋านี สิบเอ็ดจะมีจำนวนคำทั้งหมด 44 คำ จาก ทั้งหมด 8 วรรคในการทดลองวิเคราะห์กภย๋านี สิบเอ็ด โดยใช้ฐาน กฎมีร้อยละความถูกต้องเฉลี่ยของการตัดคำเท่ากับ 97.49 จากการตัด คำที่ผิดพลาดทำให้การวิเคราะห์กภย๋านี สิบเอ็ด โดยใช้ฐานกฎ ผิดพลาดไปด้วย

สรุปการนำเสนอข้างต้นเป็นการสรุปผลการวิเคราะห์ที่ได้จาก การพัฒนาขั้นตอนวิธีและกรอบแนวคิดของการวิเคราะห์ของฐานกฎ และวิธีการตัดคำที่ได้พัฒนาขึ้น นอกจากเป็นการวิเคราะห์ ฉันทลักษณ์แล้วยังเป็นการตรวจสอบขั้นตอนวิธีของฐานกฎและ วิธีการตัดคำด้วย



ภาพที่ 5: Functional Test

ภาพที่ 5 เป็นการทดสอบโปรแกรมจำลองเป็นไปตาม ความต้องการของผู้ใช้หรือไม่ ซึ่ง Efficacy คือ การประเมินด้าน ความสามารถตรงตามความต้องการของผู้ใช้ Accuracy คือ การ ประเมินด้านความถูกต้องในการทำงานของโปรแกรมจำลอง และ Suitability คือ การประเมินด้านความเหมาะสมในการใช้งาน จากการ

ประเมินผล Functional Test ทั้ง 3 ด้าน ผู้ใช้มีความพึงพอใจต่อ โปรแกรมจำลองอยู่ในระดับดีมาก

5. สรุปผลการวิจัย

การวิเคราะห์กภย๋านี สิบเอ็ด โดยใช้ฐานกฎที่ได้ทำการทดลอง มีการประเมินผลการทดลองทั้งขั้นตอนวิธีของการวิเคราะห์และ ความพึงพอใจต่อโปรแกรมจำลอง

จากการประเมินผลการทดลองที่ได้จากโปรแกรมจำลอง การวิเคราะห์กภย๋านี สิบเอ็ดสามารถวิเคราะห์ได้ถูกต้องร้อยละ 83 ความถูกต้องในการตัดคำเฉลี่ยเท่ากับร้อยละ 97.49 และการประเมินผลความพึงพอใจของผู้ใช้ต่อ โปรแกรมจำลองในการ ทดลองใช้วิเคราะห์กภย๋านี สิบเอ็ด ผู้ใช้มีความพึงพอใจอยู่ใน ระดับดีมาก ซึ่งความผิดพลาดของการวิเคราะห์ อาจเกิดจากวิธี การตัดคำที่ได้พัฒนาขึ้นหรือจำนวนคำศัพท์ในพจนานุกรมมีไม่ มากพอ จึงทำให้การตัดคำผิดและส่งผลต่อการวิเคราะห์

การพัฒนาในอนาคตเนื่องจากการวิจัยนี้เป็นเพียงการทดลองให้ ได้มาซึ่งขั้นตอนวิธีและกรอบแนวคิดของการวิเคราะห์เท่านั้นยังไม่ มีการพัฒนาเป็นระบบที่สมบูรณ์จึงสามารถนำไปพัฒนาให้สมบูรณ์ ตอบสนองความต้องการของผู้ใช้ และให้มีคามยืดหยุ่นในการ วิเคราะห์ที่มีความหลากหลายยิ่งขึ้น

เอกสารอ้างอิง

- [1] เอกรินทร์ สัมहाส และคนอื่น ๆ. ภาษาไทย ๖. พิมพ์ครั้งที่ 1. กรุงเทพฯ: อักษรเจริญทัศน์. 2552.
- [2] วรรณธ์ อักษรพงษ์ (บรรณาธิการ). ภาษาไทย 4. กรุงเทพมหานคร: สำนักพิมพ์ จุฬาลงกรณ์มหาวิทยาลัย, 2539.
- [3] ชูปณี เสงสนันธุ์. "การตัดคำภาษาไทย". วิทยานิพนธ์ปริญญาวิทยาศาสตร มหบัณฑิต สาขาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยขอนแก่น, 2549.
- [4] Paisam Charoenpomsawat, Tanja Schultz. "Improving Word Segmentation for Thai Speech Translation". 2008 : 241-244.
- [5] Choochart Haruechaiyasak, Sarawoot Kongyong and Matthew Dailey. "A Comparative Study on Thai Word Segmentation Approaches". 2008 : 125-128.
- [6] พรพิมล มุมนะวงศ์ และมณฑิธา รัตนศิริวงศ์วุฒิ. "ระบบสารสนเทศเพื่อ สนับสนุนการตัดสินใจสำหรับผู้บริหารสถานศึกษาโดยใช้เทคนิคผสมผสาน". 2553:213-218.

- [7] อนรุท จำริญนุสิทธิ์. “ระบบสารสนเทศสำหรับแบบประเมินภาระงานพนักงานของสถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ”. ปัญหาพิเศษวิทยาศาสตร์มหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2547.
- [8] Shuwei Chen, Jun Liu, Hui Wang, and Juan Carlos Augusto. “A Data Driven Rule-Base Inference Approach for Classification Systems”, 2011: 78-81.
- [9] Lorena Chavarria-Baez, Xiaou Li. “Verification of Active Rule Base via Conditional Colored Petri Nets”, 2007 : 343-348
- [10] สุจิตรา อุดลย์เกษม และคนอื่น ๆ. “การติดตามสถานะของนักศึกษาในระดับบัณฑิตศึกษาโดยใช้กราฟที่เปลี่ยนแปลงตามเวลา”. 2554: 25-31.
- [11] กิตตินันท์ ธิมาสาร. “ระบบสารสนเทศเพื่อวิเคราะห์คุณลักษณะการเลือกโปรแกรมชั้นของลูกค้ำที่ใช้บริการคงสิทธิเลขหมายโทรศัพท์เคลื่อนที่โดยใช้ฐานความรู้ร่วมกับกฎความสัมพันธ์ของข้อมูล”. ปัญหาพิเศษวิทยาศาสตร์มหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2554.
- [12] พัชรินทร์ ทองอารี. “การศึกษาเปรียบเทียบวิธีการตัดคำสำหรับการจำแนกประเภทข้อความในภาษาไทย”. วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์, 2551.

ปัจจัยที่มีอิทธิพลต่อความตั้งใจใช้ซอฟต์แวร์บริหารโครงการบนกลุ่มเมฆในองค์กรเอกชน

Factors Influencing Intention to Use Cloud-based Project Management Software in Private Organization

มนีวรรณ ศุภคิมัสโร (Maneewan Supatimassaro)¹ สุพิชญา จันทร์โพธิ์ (Supitchaya Janplong)²

และ กนกวรรณ อัจฉริยะชาญวนนิช (Kanokwan Atchariyachanvanich)³

คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
jelvy.design@gmail.com¹, nuinuy.na@gmail.com², kanokwan@it.kmitl.ac.th³

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อเสนอแบบจำลองเพื่อวิเคราะห์ปัจจัยที่มีอิทธิพลต่อความตั้งใจใช้ซอฟต์แวร์บริหารโครงการบนกลุ่มเมฆในองค์กรเอกชนในประเทศไทย โดยได้ทำการศึกษาแนวคิด ทฤษฎีรวมถึงงานวิจัยที่เกี่ยวข้องต่างๆ และได้พัฒนาเป็นแบบจำลองขึ้น โดยแบบจำลองนั้นได้นำปัจจัยจากแบบจำลองการยอมรับเทคโนโลยี TAM2 บางส่วนมาใช้ ได้แก่ การรับรู้ถึงประโยชน์ในการใช้งาน การรับรู้ถึงความง่ายในการใช้งาน และปัจจัยอีกส่วนหนึ่งเกิดจากการศึกษาวรรณกรรมที่เกี่ยวข้องและนำเสนอโดยผู้วิจัยเอง ได้แก่ การทำงานร่วมกัน การใช้งานบนอุปกรณ์เคลื่อนที่ ซึ่งแบบจำลองนี้จะช่วยให้นักพัฒนาซอฟต์แวร์ได้พัฒนาซอฟต์แวร์บริหารโครงการบนกลุ่มเมฆที่ตรงกับความต้องการของผู้ใช้มากยิ่งขึ้น โดยมีผู้ตอบแบบสอบถามจำนวน 331 คน นำข้อมูลที่ได้มาวิเคราะห์การถดถอยพหุคูณ (Multiple Regression Analysis) ซึ่งพบว่าปัจจัยที่มีอิทธิพลต่อความตั้งใจใช้ซอฟต์แวร์บริหารโครงการบนกลุ่มเมฆในองค์กรเอกชนในประเทศไทยคือ การรับรู้ถึงประโยชน์ในการใช้งาน การทำงานร่วมกัน และการใช้งานบนอุปกรณ์เคลื่อนที่

คำสำคัญ: ซอฟต์แวร์บริหารโครงการ ซอฟต์แวร์บริหารโครงการบนกลุ่มเมฆ ปัจจัยที่มีอิทธิพลต่อความตั้งใจใช้ซอฟต์แวร์

Abstract

This paper aims to investigate factors influencing intention to use cloud-based project management software in private organizations. The research model was developed based on the literature reviews and the findings from the survey of functions in the cloud-based

project management software. It consists of four factors which are Perceived Usefulness and Perceived Ease of Use from the Enhanced Technology Acceptance Model (TAM2), as well as two new factors including Collaboration and Usage on Mobile Device. The confirmatory factor analysis and multiple regression analysis were used to analyze data from 331 respondents gathered by questionnaire survey. The result shows that Perceived Usefulness, Collaboration and Usage on Mobile Device significantly affected the usage intention in cloud-based project management software. This model will be a guideline for software developers in developing the cloud-based project management software that well matches users' requirements.

Keyword: Project Management Software, Cloud-Based Project Management Software, Factors Influencing Intention to Use Software.

1. บทนำ

ในปัจจุบันการประมวลผลแบบกลุ่มเมฆ (Cloud Computing) เป็นเทคโนโลยีที่กำลังได้รับความนิยมเป็นอย่างมากทั่วโลก เนื่องจากต้นทุนที่ต่ำทำให้องค์กรขนาดเล็กสามารถนำเทคโนโลยีสารสนเทศเข้ามาใช้ได้ดียิ่งขึ้น แต่ในประเทศไทยนั้นยังไม่ค่อยได้รับความสนใจ จากการสำรวจในหัวข้อ 2013 BSA Global Cloud Computing scorecard ประเทศไทยได้ถูกจัดอันดับให้อยู่ในอันดับรองสุดท้าย และตามด้วยเวียดนามในอันดับสุดท้าย [1] รวมถึงการใช้ซอฟต์แวร์บริหารโครงการบนกลุ่มเมฆในองค์กรเอกชนในประเทศไทยอีกด้วย ซึ่งการบริหารโครงการเป็นสิ่งที่จำเป็น เนื่องจากช่วยทำให้

โครงการบรรลุเป้าหมายที่กำหนดไว้ ภายใต้ขอบเขตของงาน เวลา เงินทุน และทรัพยากรที่กำหนด เพื่อช่วยให้การบริหารโครงการมีประสิทธิภาพมากยิ่งขึ้น การเลือกใช้ซอฟต์แวร์เพื่อนำมาช่วยในการบริหารโครงการจึงเป็นสิ่งที่จำเป็น และเพื่อที่องค์กรจะได้มุ่งความสนใจไปที่ธุรกิจหลักขององค์กรเท่านั้น โดยไม่จำเป็นต้องคำนึงถึงค่าใช้จ่ายในเรื่องลิขสิทธิ์ซอฟต์แวร์ ฮาร์ดแวร์ หรือการบำรุงรักษาระบบ การเลือกใช้ซอฟต์แวร์บริหารโครงการบนกลุ่มเมฆจึงเป็นทางเลือกที่ดี เนื่องจากไม่จำเป็นต้องติดตั้งซอฟต์แวร์เพื่อใช้งาน ไม่ต้องลงทุนทางด้านฮาร์ดแวร์ และไม่ต้องสนใจในเรื่องการบำรุงรักษา

เพื่อให้องค์กรเหล่านั้นหันมาใช้ซอฟต์แวร์บริหารโครงการบนกลุ่มเมฆจึงจำเป็นที่จะต้องศึกษาหาปัจจัยที่มีอิทธิพลต่อความตั้งใจใช้ซอฟต์แวร์บริหารโครงการบนกลุ่มเมฆ เพื่อให้ นักพัฒนาซอฟต์แวร์ได้พัฒนาซอฟต์แวร์บริหารโครงการบนกลุ่มเมฆที่ตรงกับความต้องการของผู้ใช้ ดังดูให้ผู้ใช้งานหันมาสนใจและใช้งานเพิ่มมากยิ่งขึ้น

2. การทบทวนวรรณกรรม

2.1 การประมวลผลบนกลุ่มเมฆ (Cloud Computing)

จากการนิยามโดยสถาบันมาตรฐานและเทคโนโลยีแห่งชาติของสหรัฐอเมริกา (National Institute of Standards and Technology : NIST) [2] ได้ให้ความหมายว่า “การประมวลผลแบบกลุ่มเมฆเป็นแบบจำลองเพื่อความสะดวกในการใช้งาน เมื่อผู้ใช้มีความต้องการใช้งานสามารถเข้าถึงได้โดยอินเทอร์เน็ตเพื่อกำหนดค่าให้กับทรัพยากรคอมพิวเตอร์ เช่น อินเทอร์เน็ต เซิร์ฟเวอร์ แอปพลิเคชัน และการให้บริการ มีการทำงานที่รวดเร็วและมีปฏิสัมพันธ์กับผู้ให้บริการน้อย”

2.2 ซอฟต์แวร์บริหารโครงการ (Project Management Software)

ซอฟต์แวร์บริหารโครงการ คือ กลุ่มของโปรแกรมประยุกต์ที่ออกแบบมาโดยเฉพาะเพื่อช่วยบริหารจัดการ [2] ซึ่งซอฟต์แวร์บริหารโครงการสามารถแบ่งได้เป็น 2 ประเภท [3] คือ (1) On-premise มีการติดตั้งซอฟต์แวร์ในเครื่องเซิร์ฟเวอร์ขององค์กร มีข้อมูลกลางที่ถูกจัดการโดยพนักงานด้านเทคโนโลยีสารสนเทศขององค์กร (2) Online (Cloud) ซอฟต์แวร์ประเภทนี้ได้นำเทคโนโลยีการประมวลผลแบบกลุ่มเมฆเข้ามาใช้ เป็นการให้บริการแบบ software-as-a-Service (SaaS)

2.3 ซอฟต์แวร์บริหารโครงการบนกลุ่มเมฆ (Cloud-Based Project Management Software)

ซอฟต์แวร์บริหารโครงการบนกลุ่มเมฆ คือ ซอฟต์แวร์การบริหารโครงการที่นำเทคโนโลยีการประมวลผลแบบกลุ่มเมฆเข้ามาใช้ มีการให้บริการแบบ software-as-a-service (SaaS) ข้อดีของซอฟต์แวร์บริหารโครงการบนกลุ่มเมฆคือ [4] (1) เข้าถึงแฟ้มเอกสารได้ง่าย สามารถเข้าถึงข้อมูลได้ตลอดเวลาจากทุกที่ผ่านเครือข่ายอินเทอร์เน็ต (2) รับผิดชอบแบบเรียลไทม์ด้วยการเข้าถึงผ่านเครือข่ายอินเทอร์เน็ต สมาชิกในทีมสามารถรายงานสถานะการทำงาน รวมถึงอัปเดตงานที่ทำในปัจจุบันได้ (3) สามารถทำงานร่วมกันได้ง่าย เนื่องจากความสามารถในการสนทนาเป็นกลุ่มแบบออนไลน์นั้นช่วยขยับระยะเวลาในการติดต่อสื่อสารจากการส่งอีเมล

2.4 เปรียบเทียบซอฟต์แวร์บริหารโครงการแบบ On-premise และแบบ Online (Cloud)

เพื่อแสดงความแตกต่างระหว่างซอฟต์แวร์บริหารโครงการแบบ On-premise และแบบ Online จึงได้สรุปไว้ในตารางที่ 1

ตารางที่ 1: ตารางเปรียบเทียบซอฟต์แวร์บริหารโครงการแบบ On-premise และแบบ Online

	ซอฟต์แวร์บริหารโครงการแบบ On-premise	ซอฟต์แวร์บริหารโครงการแบบ Online (Cloud)
การติดตั้งซอฟต์แวร์	ซอฟต์แวร์ถูกติดตั้งอยู่ในเครื่องเซิร์ฟเวอร์ขององค์กร ต้องรอจนกระทั่งซอฟต์แวร์ติดตั้งเสร็จ จึงจะสามารถใช้งานได้	ซอฟต์แวร์ถูกติดตั้งอยู่ในเครื่องเซิร์ฟเวอร์ของผู้ให้บริการ สามารถเริ่มต้นใช้งานได้ทันทีเมื่อดำเนินการติดต่อและชำระค่าบริการกับผู้ให้บริการเรียบร้อยแล้ว
ฮาร์ดแวร์	ต้องจัดหาฮาร์ดแวร์และแพลตฟอร์มที่รองรับกับแอปพลิเคชันที่จะใช้งาน	ไม่ต้องคำนึงถึงเรื่องฮาร์ดแวร์ เนื่องจากผู้ให้บริการเป็นฝ่ายดูแลทั้งหมด

	ซอฟต์แวร์บริหารโครงการแบบ On-premise	ซอฟต์แวร์บริหารโครงการแบบ Online (Cloud)
ค่าใช้จ่าย	ค่าใช้จ่ายขึ้นอยู่กับฮาร์ดแวร์ ค่าลิขสิทธิ์ซอฟต์แวร์ ค่าใช้จ่ายในการดำเนินงาน และค่าใช้จ่ายในการบำรุงรักษา	จ่ายตามการใช้งานจริง จ่ายรายเดือน หรือจ่ายตามแพคเกจที่เลือกแล้วแต่ที่ผู้ให้บริการจะกำหนด
การขยายระบบ	หากมีจำนวนผู้ใช้ซอฟต์แวร์เพิ่มมากขึ้น อาจจำเป็นต้องเสียบค่าลิขสิทธิ์ซอฟต์แวร์เพิ่ม รวมถึงการจัดหาฮาร์ดแวร์เพื่อรองรับการทำงานที่หนักขึ้น	หากมีจำนวนผู้ใช้ซอฟต์แวร์เพิ่มมากขึ้น สามารถดำเนินการติดต่อกับผู้ให้บริการเพื่อเริ่มต้นการใช้งานได้ทันที
การเข้าถึงซอฟต์แวร์	เข้าถึงผ่านระบบเครือข่ายภายในขององค์กร	เข้าถึงผ่านทางเครือข่ายอินเทอร์เน็ตหรือผ่านแอปพลิเคชันบนโทรศัพท์มือถือ (ขึ้นอยู่กับผู้ให้บริการ)
ความปลอดภัยของข้อมูล	มีความเสี่ยงน้อยเนื่องจากข้อมูลถูกจัดเก็บอยู่ในเครื่องเซิร์ฟเวอร์ขององค์กร แต่หากเครื่องเซิร์ฟเวอร์มีปัญหา ข้อมูลเหล่านั้นอาจสูญหายได้	มีความเสี่ยงน้อยเนื่องจากการดูแลรักษาความปลอดภัยโดยผู้ให้บริการ หากเครื่องเซิร์ฟเวอร์ของผู้ให้บริการมีปัญหา ก็สามารถมั่นใจได้ว่าข้อมูลที่ถูเก็บไว้จะไม่หาย เนื่องจากใช้ระบบ Cloud Computing

2.5 แบบจำลองการยอมรับเทคโนโลยี (Enhanced Technology Acceptance Model: TAM2)

แบบจำลอง TAM2 นำเสนอโดย Venkatesh และ Davis [5] ถูกพัฒนามาจากแบบจำลอง TAM เพื่อช่วยให้สามารถพยากรณ์พฤติกรรมการใช้งานเทคโนโลยีสารสนเทศได้ชัดเจนยิ่งขึ้น โดยได้รับการปรับปรุงในส่วนของตัวแปรภายนอกและปัจจัยที่เกิดก่อน โดยสามารถแบ่งได้เป็น 2 กระบวนการ [6,7] คือ (1) กระบวนการของอิทธิพลจากสังคม (Social influence process) ได้แก่ บรรทัดฐานของบุคคลที่อยู่โดยรอบการแสดงพฤติกรรม (Subjective norm) ความสมัครใจ (Voluntariness) และภาพลักษณ์ (Image) (2) กระบวนการใช้ปัญญา (Cognitive instrumental process) ได้แก่ ความเกี่ยวข้องสัมพันธ์กับงาน (Job relevance) คุณภาพของผลลัพธ์ (Output quality) ผลลัพธ์ที่สามารถแสดงให้เห็นก่อนได้ (Results demonstrability) และการรับรู้ว่าเป็นระบบที่ง่ายต่อการใช้งาน (Perceived ease of use)

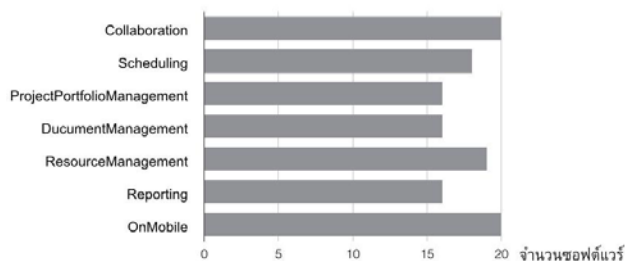
3. ขั้นตอนวิธีดำเนินงานวิจัย

3.1 การกำหนดปัจจัย

จากการเก็บรวบรวมข้อมูลที่ได้จากการทบทวนวรรณกรรม

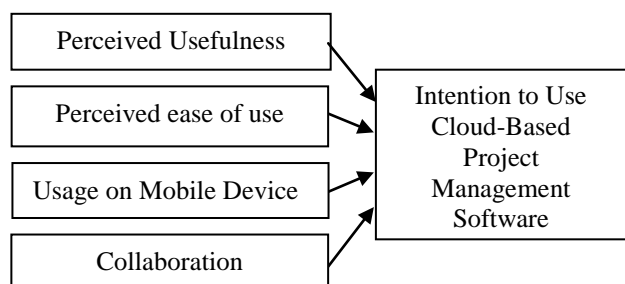
บทความและทฤษฎีต่างๆ ผู้ศึกษาจึงได้มีการพัฒนากรอบแนวคิดซึ่งมีต้นแบบมาจากทฤษฎีการยอมรับเทคโนโลยีได้ดังนี้ (1) ปัจจัยด้านการรับรู้ถึงประโยชน์ในการใช้งาน (Perceived Usefulness) (2) ปัจจัยด้านการรับรู้ถึงความง่ายในการใช้งาน (Perceived Ease of Use)

จากการสำรวจฟังก์ชันการทำงานของซอฟต์แวร์บริหารโครงการในกลุ่มเมฆจำนวน 20 ซอฟต์แวร์ ที่เปิดให้บริการอยู่ในปัจจุบัน ซึ่งเป็นซอฟต์แวร์ที่ให้บริการแบบ Public Cloud Computing เช่น AffinityLive, Clarizen, Comindwork สามารถสรุปได้ดังแสดงในภาพที่ 1



ภาพที่ 1 : ผลสำรวจฟังก์ชันของซอฟต์แวร์บริหารโครงการในกลุ่มเมฆ

จะเห็นได้ว่าฟังก์ชันที่ทุกซอฟต์แวร์มีเหมือนกันมากที่สุด นั่นก็คือ การทำงานร่วมกันและการใช้งานบนอุปกรณ์เคลื่อนที่ ซึ่งทั้งสองฟังก์ชันนี้เป็นฟังก์ชันที่น่าสนใจ ถึงแม้ซอฟต์แวร์บริหารโครงการแบบ On-premise จะสามารถทำงานร่วมกันได้เช่นเดียวกับซอฟต์แวร์บริหารโครงการบนกลุ่มเมฆ แต่เป็นการทำงานร่วมกันภายในองค์กรเท่านั้น และการทำงานบนอุปกรณ์เคลื่อนที่ที่เป็นฟังก์ชันที่มีเฉพาะซอฟต์แวร์บริหารโครงการบนกลุ่มเมฆ ในส่วนของฟังก์ชันอื่นๆ นั้นจะสามารถพบได้ในซอฟต์แวร์ทั้งสองประเภท ผู้วิจัยจึงได้เสนอปัจจัยใหม่ขึ้นได้แก่ (1) ปัจจัยด้านการทำงานร่วมกัน (Collaboration) (2) ปัจจัยด้านการใช้งานบนอุปกรณ์เคลื่อนที่ (Usage on Mobile Device) โดยสามารถสรุปกรอบงานวิจัยได้ดังภาพที่ 2



ภาพที่ 2 : กรอบการวิจัย (Research Model)

สมมติฐานตามกรอบงานวิจัยประกอบไปด้วย

สมมติฐานที่ 1 การรับรู้ถึงประโยชน์ในการใช้งานมีผลกระทบทางบวกต่อความตั้งใจในการใช้ซอฟต์แวร์บริหารโครงการบนกลุ่มเมฆ

สมมติฐานที่ 2 การรับรู้ถึงความง่ายในการใช้งานมีผลกระทบทางบวกต่อความตั้งใจในการใช้ซอฟต์แวร์บริหารโครงการบนกลุ่มเมฆ

สมมติฐานที่ 3 คุณสมบัตินี้ด้านการทำงานร่วมกันมีผลกระทบทางบวกต่อความตั้งใจในการใช้ซอฟต์แวร์บริหารโครงการบนกลุ่มเมฆ

สมมติฐานที่ 4 คุณสมบัตินี้ด้านการใช้งานบนอุปกรณ์เคลื่อนที่มีผลกระทบทางบวกต่อความตั้งใจในการใช้ซอฟต์แวร์บริหารโครงการบนกลุ่มเมฆ

3.2 ประชากรและกลุ่มตัวอย่าง

ประชากรที่ใช้ในการวิจัยนี้คือ กลุ่มของผู้มีส่วนร่วมในการบริหารโครงการและผู้มีอำนาจในการตัดสินใจซอฟต์แวร์ที่เคยใช้ซอฟต์แวร์บริหารโครงการซึ่งทำงานในองค์กรเอกชนในประเทศไทย ตำแหน่งงานที่เกี่ยวข้องกับการพัฒนาซอฟต์แวร์

การเลือกกลุ่มตัวอย่างใช้วิธีการเลือกแบบไม่ใช้หลักความน่าจะเป็น (Nonprobability Sampling) ซึ่งจะเลือกกลุ่มตัวอย่างแบบเจาะจง (Purposive Sampling) และแบบลูกโซ่ (Snowball Sampling) ซึ่งมีลักษณะตรงกับกลุ่มประชากรที่ได้ตั้งไว้ สามารถคำนวณหาขนาดของกลุ่มตัวอย่าง โดยใช้แนวคิดของ Taro Yamane ซึ่งสามารถคำนวณได้จากสูตร [8]

$$n = \frac{N}{1+N(e)^2} \tag{1}$$

เมื่อ n คือ ขนาดของกลุ่มตัวอย่าง

N คือ ขนาดของกลุ่มประชากร

e คือ ความคาดเคลื่อนของการสุ่มตัวอย่าง

คำนวณที่ระดับความเชื่อมั่น 95% ค่าความคาดเคลื่อน 5% ได้ขนาดกลุ่มตัวอย่างจำนวน 393 คน และมีการตอบแบบสอบถามกลับมาจำนวน 331 คน คิดเป็น 84.22%

3.3 เครื่องมือที่ใช้ในการวิจัยและการตรวจสอบความเชื่อถือได้ของเครื่องมือ

เครื่องมือที่ใช้ในการวิจัย ใช้เพื่อเก็บรวบรวมข้อมูลซึ่งเป็นแบบสอบถาม (Questionnaire) ที่ผู้วิจัยได้พัฒนาขึ้นเองโดยการนำแนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้องมาปรับปรุงให้เหมาะสม โดยแบบสอบถามแบ่งออกเป็นสองส่วนคือ (1) แบบเติมคำ (Fill-in-the-blank) จำนวน 4 ข้อ (2) แบบมาตราประมาณค่า (Rating Scale) 5 ระดับตามวิธีของลิเคิร์ต (Likert) จำนวน 19 ข้อ

การวัดความตรงของเนื้อหา (Content Validity) ผู้วิจัยได้นำแบบสอบถามเสนอต่อผู้เชี่ยวชาญจำนวน 3 ท่าน เพื่อทำการประเมิน หาค่าสัมประสิทธิ์ความสอดคล้อง (Index of Item Objective Congruence: IOC) และตรวจสอบความถูกต้องของเนื้อหา ซึ่งสามารถคำนวณได้จากสูตร [9]

$$IOC = \frac{\sum R}{N} \tag{2}$$

เมื่อ IOC คือ ค่าสัมประสิทธิ์ความสอดคล้อง

ΣR คือ ผลรวมของคะแนนที่ได้จากผู้เชี่ยวชาญทั้งหมด

N คือ จำนวนผู้เชี่ยวชาญ

การวัดความเที่ยงตรงตามโครงสร้าง (Construct Validity) ผู้วิจัยได้เลือกใช้การวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Analysis: CFA) ซึ่งเป็นการวิเคราะห์ปัจจัยที่มีการทราบถึงความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตาม พบว่าค่า Kaiser Meyer Olkin (KMO) ซึ่งสามารถคำนวณได้จากสูตร [10]

$$KMO = \frac{\sum r_i^2}{\sum r_i^2 + \sum (\text{partial correlation})^2} \quad (3)$$

เมื่อ r คือ สัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปร ซึ่งทำให้ค่า $0 \leq KMO \leq 1$

มีค่าเท่ากับ 0.829 ซึ่งมากกว่าค่ามาตรฐานที่ 0.50 แสดงให้เห็นว่าข้อมูลที่มีอยู่นั้นเหมาะสมกับการทำ Factor Analysis และ Bartlett's Test มีนัยสำคัญทางสถิติเท่ากับ 0.000 ซึ่งค่ามาตรฐานกำหนดไว้ที่น้อยกว่า 0.05 (Sig. < 0.05) แสดงให้เห็นว่าตัวแปรมีความสัมพันธ์กัน เหมาะสมที่จะใช้วิธีวิเคราะห์องค์ประกอบ

การวัดความเชื่อถือได้ภายใน (Internal Consistency) ผู้วิจัยได้ทดสอบแบบสอบถามจากกลุ่มตัวอย่าง 30 ชุด เพื่อหาความเชื่อถือได้ภายในของแบบสอบถาม โดยการหาค่าสัมประสิทธิ์อัลฟาของครอนบาค (Cronbach's alpha Coefficient) ซึ่งมีสูตรการคำนวณดังนี้ [10]

$$\alpha = \frac{k\bar{r}}{1+\bar{r}(k-1)} \quad (4)$$

เมื่อ α คือ ค่าสัมประสิทธิ์อัลฟาของครอนบาค

k คือ จำนวนคำถาม

\bar{r} คือ ค่าเฉลี่ยของค่าสัมประสิทธิ์สหสัมพันธ์ของ

ทุกคำถาม

จากคำถามทั้งหมด 19 ข้อ มีค่าสัมประสิทธิ์อัลฟาของครอนบาคเท่ากับ 0.883 ซึ่งมากกว่าค่ามาตรฐานที่ 0.70 แสดงว่าแบบสอบถามนี้มีความเชื่อถือได้

3.4 การวิเคราะห์ข้อมูล

ผู้วิจัยได้นำแบบสอบถามที่มีความครบถ้วนสมบูรณ์ของคำตอบมาประมวลผลโดยใช้โปรแกรมการวิเคราะห์สถิติด้วยคอมพิวเตอร์ โดยแบ่งออกเป็น 2 ส่วน (1) สถิติเชิงพรรณนา (Descriptive Statistic) ใช้เพื่อสรุปสาระสำคัญของข้อมูลที่เก็บรวบรวมมาได้ โดยแสดงอยู่ในรูปแบบค่าสถิติได้แก่ ค่าเฉลี่ย (Mean) และค่าเบี่ยงเบนมาตรฐาน (Standard Deviation) (2) สถิติเชิงอนุมาน (Inferential Statistics) ใช้เพื่อพิสูจน์และสรุปสมมุติฐานที่ได้ตั้งขึ้น โดยการวิเคราะห์การถดถอยพหุคูณ (Multiple Regression Analysis) ซึ่งเป็นวิธีการที่ใช้เพื่อศึกษาหาความสัมพันธ์ของตัวแปร 2 ประเภทคือ ตัวแปรอิสระมากกว่า 1 ตัวแปร กับตัวแปรตาม 1 ตัวแปร

4. ผลการดำเนินการวิจัย

4.1 การวิเคราะห์สถิติเชิงพรรณนา (Descriptive Statistics)

จากผู้ตอบแบบสอบถามจำนวน 331 คน คิดเป็น 100% สามารถสรุปข้อมูลได้ว่า ผู้ตอบแบบสอบถาม 71.30% ใช้ซอฟต์แวร์บริหารโครงการบนกลุ่มเมฆ มีประสบการณ์ในการใช้ซอฟต์แวร์บริหารโครงการมาแล้ว 1-2 ปี คิดเป็น 46.10% มีผู้มีส่วนร่วมในการบริหารโครงการ 1-5 คน คิดเป็น 40.60% และมีงบประมาณในการจัดซื้อซอฟต์แวร์บริหารโครงการน้อยกว่า 10,000 บาท คิดเป็น 23.90% และสรุประดับความคิดเห็นของปัจจัยที่มีอิทธิพลต่อความตั้งใจใช้ซอฟต์แวร์บริหารโครงการบนกลุ่มเมฆในองค์กรเอกชนแสดงดังตารางที่ 2

ตารางที่ 2: ระดับความคิดเห็นของปัจจัยที่มีอิทธิพลต่อความตั้งใจใช้ซอฟต์แวร์บริหารโครงการบนกลุ่มเมฆในองค์กรเอกชน

ปัจจัยที่มีอิทธิพลต่อความตั้งใจใช้ซอฟต์แวร์บริหารโครงการบนกลุ่มเมฆในองค์กรเอกชน	ค่าเฉลี่ย (Mean)	ค่าเบี่ยงเบนมาตรฐาน (S.D)	ระดับความคิดเห็น
ด้านการรับรู้ถึงประโยชน์ในการใช้งาน	4.41	0.789	เห็นด้วยมาก
ด้านการรับรู้ถึงความง่ายในการใช้งาน	4.26	0.758	เห็นด้วยมาก
ด้านการทำงานร่วมกัน	4.52	0.679	เห็นด้วยมากที่สุด
ด้านการใช้งานบนอุปกรณ์เคลื่อนที่	4.49	0.7274	เห็นด้วยมาก
ด้านความตั้งใจใช้ซอฟต์แวร์บริหารโครงการบนกลุ่มเมฆ	3.98	0.863	เห็นด้วย

4.2 สถิติเชิงอนุมาน (Inferential Statistics)

สามารถแสดงผลการวิเคราะห์การถดถอยพหุคูณ ได้ดัง

ตารางที่ 3

ตารางที่ 3: ผลการวิเคราะห์การถดถอยพหุคูณ

Model	B	Std. Error	Beta	t	Sig.
(Constant)	-1.001E-013	.053		.000	1.000
PU	.181	.053	.181	3.387	.001
CO	.162	.053	.162	3.045	.003
MD	.108	.053	.108	2.035	.043

$$R = .417, R^2 = .174, R^2_{adj} = .166$$

หมายเหตุ ทุกปัจจัยมีนัยสำคัญทางสถิติที่ระดับ 0.01

ในงานวิจัยนี้เลือกใช้ค่า R^2_{adj} คือค่าสัมประสิทธิ์การตัดสินใจที่ถูกปรับค่าแล้ว โดยค่าสัมประสิทธิ์การตัดสินใจพิจารณาจากความผันแปรของ y มีผลมาจาก x มากน้อยเพียงใด มีค่าตั้งแต่ 0 ถึง 1 หากค่านี้มีค่าใกล้เคียงกับ 1 มาก แสดงว่าตัวแปร

5. สรุป

5.1 สรุปผลการวิจัย

ปัจจัยที่มีอิทธิพลต่อความตั้งใจใช้ซอฟต์แวร์บริหารโครงการบนกลุ่มเมฆในองค์กรคือ การรับรู้ถึงประโยชน์ในการใช้งาน การทำงานร่วมกัน และการใช้งานบนอุปกรณ์เคลื่อนที่ ซึ่งสนับสนุนสมมุติฐานข้อที่ 1, 3 และ 4

5.2 ข้อเสนอแนะ

เนื่องจากค่า Adjusted R Square มีค่าค่อนข้างน้อย เพื่อช่วยให้ค่านี้สูงขึ้น ผู้วิจัยเสนอแนะว่าควรเพิ่มปัจจัยอื่นๆ ที่เกี่ยวข้องกับองค์กร ซึ่งอาจมีอิทธิพลต่อความตั้งใจใช้ซอฟต์แวร์บริหารโครงการบนกลุ่มเมฆ เช่น ขนาดองค์กร งบประมาณ

5.3 ข้อจำกัดของงานวิจัย

จำนวนของผู้ตอบแบบสอบถามที่มีเพียง 331 คน ซึ่งน้อยกว่าจำนวนของกลุ่มตัวอย่างที่คำนวณได้คือ 393 คน

5.4 แนวทางการวิจัยในอนาคต

เพิ่มปัจจัยที่เกี่ยวข้องกับองค์กรเพื่อหาอิทธิพลต่อความตั้งใจใช้ซอฟต์แวร์บริหารโครงการบนกลุ่มเมฆ

เอกสารอ้างอิง

[1] BSA. "2013 BSA Global Cloud Computing scorecard". Retrieved from: <http://cloudscorecard.bsa.org/2013/>

[2] P. Mell, T. Grance. "The NIST Definition of Cloud Computing", *Technical report*, National Institute of Standards and Technology. 2009.

[3] GFI Software. "ON-PREMISE VS. CLOUD-BASED SOLUTIONS". Retrieved from: http://www.gfi.com/whitepapers/Hybrid_Technology.pdf

[4] Jason, W. "3 Benefits Of Online Project Management Software" Retrieved from: <http://project-management.com/3-benefits-of-online-project-management-software/>

[5] Venkatesh, V. and Davis, F.D. "A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies," *Management Science*, 46, 2000, 186-204.

[6] สิงหะ จวีสุข และสุนันทา วงศ์จตุรภัทร. "ทฤษฎีการยอมรับการใช้เทคโนโลยีสารสนเทศ". *วารสารเทคโนโลยีสารสนเทศภาคการวิจัย*, 1(1).

[7] ปราโมทย์ ลีอนาม. "แนวความคิด และวิวัฒนาการของแบบจำลองการยอมรับการใช้เทคโนโลยี". *วารสารการจัดการสมัยใหม่*, ปีที่ 9 ฉบับที่ 1

[8] Yamane, Taro. *Statistics, An Introductory Analysis*, 2nd Ed., New York: Harper and Row. 1967.

[9] เขวาร์ตน์ เตมียกูล. เครื่องมือที่ใช้ในการวิจัย. ค้นเมื่อ 20 มกราคม 2557, จาก <http://www.kruchaow.com/kr-research/renew6.pdf>

[10] กัลยา วานิชย์บัญชา. *การวิเคราะห์สถิติ: สถิติสำหรับการบริหารและวิจัย*. กรุงเทพมหานคร : โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย. 2555.

การเปรียบเทียบความพยายามการสร้างกรณีทดสอบในระดับการทดสอบยูนิต สำหรับซอฟต์แวร์ที่ประยุกต์ใช้ดีไซน์แพตเทิร์น

ระหว่างวิธีการสร้างกรณีทดสอบแบบปกติ และวิธีการรีแฟกทอริง

A Comparison Study of Test Effort between Manually Revised Test Case and Test Case Refactoring for Pattern-Based Software Development

จตุรดา ดีอิ่ง (Jaturada Deeying)¹ และอัสฎาพร ทรัพย์สมบูรณ์ (Assadaporn Sapsomboon)²

¹นิสิตปริญญาโท สาขาวิชาการพัฒนาซอฟต์แวร์ทางด้านธุรกิจ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย

²รองศาสตราจารย์ ประจำภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย

Jaturada.deeying@gmail.com¹, assadaporn@cbs.chula.ac.th²

บทคัดย่อ

การทดสอบระดับยูนิตเป็นการทดสอบระดับย่อย ดำเนินการทดสอบด้วยกรณีทดสอบเพื่อยืนยันการทำงานของซอฟต์แวร์ถึงพฤติกรรมการทำงานที่สอดคล้องตรงความต้องการ แต่ซอฟต์แวร์ภายหลังส่งมอบอาจมีการเปลี่ยนแปลงเพื่อปรับปรุงคุณภาพ ผู้ทดสอบจึงต้องสร้างกรณีทดสอบใหม่ให้สอดคล้องกับซอฟต์แวร์ที่เปลี่ยนแปลงเพื่อใช้ทดสอบผลการทำงานที่คงเดิม เนื่องจากการสร้างกรณีทดสอบเป็นขั้นตอนที่ต้องใช้ความพยายามมากและมีความซับซ้อน และได้มีนักวิจัยเสนอหลักการรีแฟกทอริงกรณีทดสอบเพื่อปรับปรุงกรณีทดสอบเพียงบางส่วนที่ได้รับผลกระทบ ดังนั้นจึงเป็นประเด็นที่น่าสนใจที่จะวิเคราะห์ถึงวิธีการสร้างกรณีทดสอบ ดังนี้ (1) วิธีสร้างกรณีทดสอบตามปกติ (2) วิธีสร้างกรณีทดสอบด้วยการรีแฟกทอริง ที่ส่งผลต่อความพยายามและความถูกต้องในการสร้างกรณีทดสอบ

ผลการการศึกษาแสดงว่าการสร้างกรณีทดสอบด้วยวิธีสร้างกรณีทดสอบตามปกติ และวิธีสร้างกรณีทดสอบด้วยการรีแฟกทอริง มีผลต่อความพยายาม แต่ไม่มีผลต่อความถูกต้องข้อสรุปที่ได้จากการศึกษานี้ช่วยให้ผู้สร้างกรณีทดสอบใช้เป็นข้อมูลในการเลือกวิธีการสร้างกรณีทดสอบที่เหมาะสมต่อการใช้งาน

คำสำคัญ: กรณีทดสอบ รีแฟกทอริงกรณีทดสอบ ดีไซน์แพตเทิร์น

Abstract

Unit testing is a kind of testing using test cases which verifies the functionality or behavior of the software meeting the requirements. Nevertheless, software can be changed after delivery in order to improve the quality. This hinders the ability of the original set of test cases. Tester has to revise test cases in response to the original requirements of the changed software. This process is complicated and time-consuming. Recently, researchers proposed test case refactoring approach for pattern-based refactoring. Therefore, this research study the test effort between manually revised test case and test case refactoring.

The result shows that revising test cases manually and refactoring affects the test effort, but not the correctness of test cases.

Keywords: Test Case, Test Case Refactoring, Design Pattern

1. บทนำ

ในปัจจุบัน การเปลี่ยนแปลงซอฟต์แวร์อาจเกิดขึ้นได้ภายหลังส่งมอบ เพื่อแก้ไขข้อผิดพลาด ปรับปรุงประสิทธิภาพ หรือปรับปรุงซอฟต์แวร์ให้เหมาะสมกับสภาพแวดล้อมที่เปลี่ยนไป [1] การเปลี่ยนแปลงซอฟต์แวร์ในที่นี้เรียกว่าการรีแฟกทอริง ซึ่งเป็นการปรับปรุงซอร์สโค้ดโดยไม่เปลี่ยนแปลงความหมาย (Semantic) หรือพฤติกรรมการทำงานภายนอก

(Behavior) [2][3] การรีแฟกทอริงมีจุดมุ่งหมายเพื่อแก้ไข ปัญหาหรือป้องกันปัญหาที่อาจเกิดขึ้นในอนาคต [2][4] การรีแฟกทอริงซอฟต์แวร์อาจเกิดจากการประยุกต์ใช้ดีไซน์แพตเทิร์น [5] เนื่องจากดีไซน์แพตเทิร์นเป็นแบบแผนแนวทางที่ใช้แก้ไขหรือเพิ่มความยืดหยุ่นในการออกแบบซอฟต์แวร์ ช่วยปรับปรุงให้ซอร์สโค้ด อ่านได้ง่าย (Readability) ยืดหยุ่นต่อความเปลี่ยนแปลง (Flexibility) และขยายโครงสร้างได้ง่าย (Extendibility) [6]

ในการยืนยันการทำงานของซอฟต์แวร์ว่าสามารถทำงานได้ตรงตามความต้องการของผู้ใช้ก่อนการส่งมอบซอฟต์แวร์จะต้องมีการทดสอบ [7] การปรับปรุงซอฟต์แวร์หรือการรีแฟกทอริงโดยประยุกต์ดีไซน์แพตเทิร์นภายหลังการส่งมอบนั้นอาจส่งผลกระทบต่อกรณีทดสอบ มีผลทำให้กรณีทดสอบเดิมไม่สามารถใช้ยืนยันการทำงานของซอฟต์แวร์ได้อีก จึงต้องมีการปรับปรุงกรณีทดสอบให้สอดคล้องต่อซอร์สโค้ดที่เปลี่ยนแปลง ในปัจจุบันมีเฟรมเวิร์กสำหรับช่วยสร้างกรณีทดสอบเข้ามาช่วยลดความพยายามของการสร้างกรณีทดสอบใหม่ โดยมีการเสนอหลักการปรับปรุงกรณีทดสอบภายหลังการปรับปรุงซอฟต์แวร์เฉพาะส่วนที่ได้รับผลกระทบ เพื่อช่วยลดความพยายามของการสร้างกรณีทดสอบใหม่ทั้งหมด

ผู้วิจัยจึงต้องการศึกษาถึงความพยายามและความถูกต้องของกรณีทดสอบจากการปรับปรุงกรณีทดสอบในวิธีที่ต่างกััน ดังนั้นวัตถุประสงค์ของการศึกษานี้ คือ (1) เพื่อเปรียบเทียบความพยายามจากวิธีการสร้างกรณีทดสอบ ระหว่างวิธีสร้างกรณีทดสอบใหม่ และวิธีสร้างกรณีทดสอบด้วยการรีแฟกทอริง (2) เพื่อเปรียบเทียบความถูกต้องของกรณีทดสอบจากวิธีสร้างกรณีทดสอบใหม่ และวิธีสร้างกรณีทดสอบด้วยการรีแฟกทอริง

2. วรรณกรรมที่เกี่ยวข้อง

2.1 ดีไซน์แพตเทิร์น

ดีไซน์แพตเทิร์นแบ่งตามวัตถุประสงค์ของการใช้งานได้ 3 ประเภท คือ 1) แพตเทิร์นการสร้างอ็อบเจกต์ (Creation patterns) เช่น เดคอเรเตอร์แพตเทิร์น 2) แพตเทิร์นโครงสร้าง (Structural patterns) เช่น ออบเจิร์ฟเวอร์แพตเทิร์น และ 3)

แพตเทิร์นพฤติกรรม (Behavioral patterns) เช่น แอบสแตรกแพตเทิร์น

2.2 การทดสอบระดับยูนิต

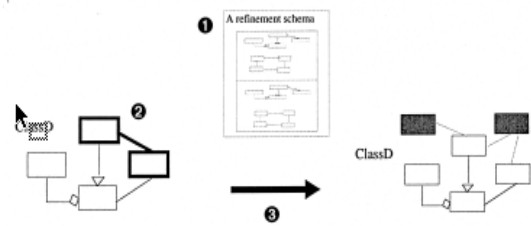
การทดสอบระดับยูนิต คือการทดสอบเพื่อยืนยันความถูกต้องการทำงานเชิงฟังก์ชันของส่วนเฉพาะหนึ่งๆภายในซอร์สโค้ด สำหรับการพัฒนาซอฟต์แวร์เชิงวัตถุ หนึ่งยูนิต หมายถึงคลาส [7] การทดสอบระดับยูนิตจัดเป็นการทดสอบประเภทไวท์บ็อกซ์ ดำเนินการทดสอบด้วยกรณีทดสอบซึ่งกำหนด ค่าทดสอบ (Input) และ ค่าคาดหวัง (Expect) เพื่อยืนยัน (Assert) ผลลัพธ์ (output) จากการทำงานของโปรแกรม เปรียบเทียบกับค่าคาดหวัง การทดสอบระดับยูนิตเพียงอย่างเดียวไม่สามารถยืนยันการทำงานร่วมกันทั้งซอฟต์แวร์ได้ แต่สามารถให้ความมั่นใจได้ว่าในแต่ละยูนิตภายในซอฟต์แวร์นั้นสามารถทำงานได้อย่างถูกต้อง เมื่อทำงานแยกเป็นอิสระจากส่วนอื่นๆ

2.3 ความครอบคลุมของกรณีทดสอบ

การทดสอบความครอบคลุมของกรณีทดสอบ คือการค้นหาบริเวณของซอร์สโค้ดที่ไม่ถูกดำเนินการทดสอบด้วยชุดของกรณีทดสอบ เป็นการวัดเชิงปริมาณจากจำนวนบรรทัดคำสั่งในซอร์สโค้ดที่ถูกดำเนินการ ความครอบคลุมของกรณีทดสอบช่วยสะท้อนคุณภาพซอฟต์แวร์ทางอ้อมได้ [8] งานวิจัยในอดีตได้มีค้นพบว่าความครอบคลุมของกรณีทดสอบควรรอยู่ที่ 80-90% [5][9]

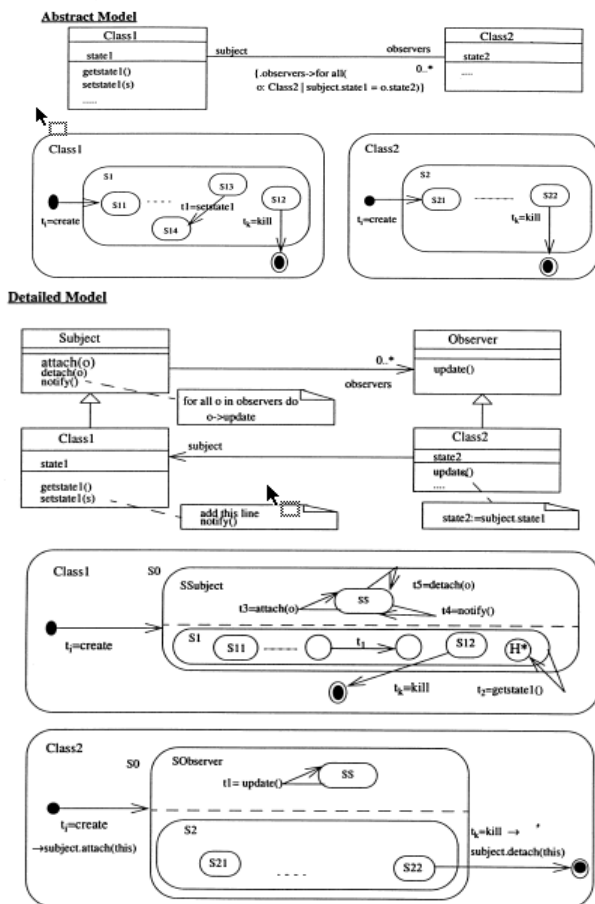
2.4 แบบแผนขั้นตอนการประยุกต์ดีไซน์แพตเทิร์น

งานวิจัยในอดีตได้เสนอหลักการวิเคราะห์ขั้นตอนการประยุกต์ดีไซน์แพตเทิร์น เพื่อสร้างแบบแผนขั้นตอนการประยุกต์ดีไซน์แพตเทิร์น เพื่อใช้สำหรับสร้างแบบแผนการแปลงการออกแบบระดับบน (High-level Design) สู่ออกแบบระดับล่าง (Low-level Design) [10] โดยขั้นตอนการแปลงการออกแบบ ดำเนินการบนพื้นฐานของการประยุกต์ดีไซน์แพตเทิร์น ภาพที่ 1 แสดงขั้นตอนการแปลงการออกแบบ ดังนี้ (1) วิเคราะห์แผนภาพ (แผนภาพคลาส แผนภาพสเตท) เพื่อเลือกโครงสร้างการปรับปรุง (Refinement Schema) ที่เหมาะสม (2) ระบุงอค์ประกอบในแผนภาพที่ต้องปรับปรุง (3) ปรับปรุงแผนภาพตามขั้นตอนโครงสร้าง



ภาพที่ 1: ขั้นตอนการแปลง [10]

โครงสร้างการปรับปรุง (Refinement Schema) แสดงได้ ด้วยแผนภาพดังนี้ (1) แผนภาพอธิบายแบบจำลองแอบสแตรก (Abstract model) (2) แผนภาพอธิบายแบบจำลองรายละเอียด หลังประยุกต์ดีไซน์แพทเทิร์น (Detail model) ดังภาพที่ 2



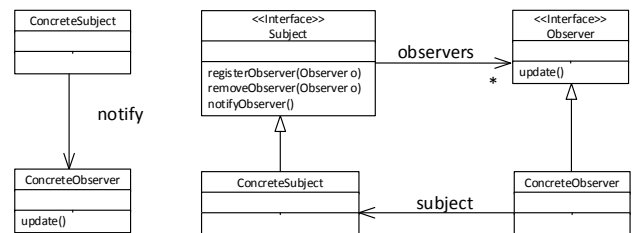
ภาพที่ 2: โครงสร้างการปรับปรุงของออบเซิร์ฟเวอร์แพทเทิร์น[10]

ภาพที่ 2 แสดงตัวอย่างโครงสร้างการปรับปรุงของออบเซิร์ฟเวอร์แพทเทิร์น ประกอบด้วยลำดับการปรับปรุงย่อย (Micro-refinement) 4 ลำดับคือ (1) การสืบทอด (Inheritance) (2) การเพิ่มการกระทำเพื่อการเปลี่ยนแปลง (Adding an action in a transition) (3) การเปลี่ยนแปลงความสัมพันธ์ (Association change) (4) การแจ้งการเปลี่ยนแปลงอัตโนมัติ (Automating notification)

2.5 การรีแฟกทอริงกรณีทดสอบ

การรีแฟกทอริงกรณีทดสอบ คือ การปรับปรุงกรณีทดสอบเดิม ซึ่งในปัจจุบันมีการศึกษาพบว่า การรีแฟกทอริงซอฟต์แวร์โดยประยุกต์ใช้ดีไซน์แพทเทิร์นนั้น มีขั้นตอนที่มีความสัมพันธ์ต่อการรีแฟกทอริงกรณีทดสอบ มิงงานวิจัยในอดีตได้ศึกษาขั้นตอนการรีแฟกทอริงกรณีทดสอบสร้างได้จากขั้นตอนการประยุกต์ดีไซน์แพทเทิร์น [5] ซึ่งมีรายละเอียดที่ขั้นตอนดังนี้

1. จำลองโครงสร้างแพทเทิร์น (Modeling pattern structure) เพื่อวิเคราะห์โครงสร้างคลาสก่อนและหลังประยุกต์ดีไซน์แพทเทิร์นบนนิยามพื้นฐานของแพทเทิร์น ดังภาพที่ 3



ภาพที่ 3: โครงสร้างคลาสก่อนและหลังประยุกต์ออบเซิร์ฟเวอร์

แพทเทิร์น

2. กำหนดการกระทำพื้นฐานในการรีแฟกทอริง (Identifying generic pattern refactoring) จากขั้นตอนที่ 1 เพื่อหาขั้นตอนรีแฟกทอริงแพทเทิร์น เช่น การสืบทอด (inheritance) การเปลี่ยนแปลงความสัมพันธ์ (association change) เป็นต้น ได้จากการประยุกต์โครงสร้างการปรับปรุงจากหัวข้อที่ 2.4 และได้ผลลัพธ์ดังภาพที่ 4 แสดงตัวอย่างขั้นตอนการสืบทอดจากการประยุกต์แพทเทิร์น

ID	Description
gpr2	gpr2.1: Add an abstract class (or interface) <<AbstractClass >> as the super class of <<ConcreteClass>> gpr2.2: Let <<ConcreteClass>> extend (or implement) << AbstractClass>>

ภาพที่ 4: การกระทำพื้นฐานในการรีแฟกทอริง[5]

3. กำหนดขั้นตอนการรีแฟกทอริงโดยเฉพาะ (Defining specific pattern refactoring) เพื่อพิจารณาแยกย่อยขั้นตอนการกระทำพื้นฐานจากขั้นตอนที่ 2 จากภาพที่ 5 แสดงตัวอย่างขั้นตอนการสืบทอดอย่างละเอียด โดยเฉพาะ

ID	Description
spr1 (gpr)	Add an interface <<Subject>> as the super class of <<ConcreteSubject >> Let <<ConcreteSubject>> implement <<Subject >>
spr2 (gpr2)	Add an interface <<Observer>> as the super class of <<ConcreteObserver>> which includes an abstract method update() Let <<ConcreteObserver>> implement <<Observer>>

ภาพที่ 5: การรีแฟกทอริงโดยเฉพาะ[5]

4. รีแฟกทอริงกรณีทดสอบ (Constructing test case refactoring) เพื่อนำขั้นตอนการรีแฟกทอริงแพดเทิร์นไปปรับปรุงกรณีทดสอบให้สอดคล้องต่อความเปลี่ยนแปลงที่เกิดขึ้นบนซอร์สโค้ด ยกตัวอย่างเช่น การสืบทอด ในขั้นตอนนี้ต้องระบุว่า เพิ่มคลาสอะไร และเกิดการเปลี่ยนแปลงอะไรที่กรณีทดสอบ ดังภาพที่ 6

Specific pattern refactorings	Test case refactorings
spr1: add Subject interface 1. Create an interface Subject 2. Let all subjects implement Subject	tr1: add test for Subject 1. Create abstract test case SubjectTest with methods makeConcreteSubject() and setUp() 2. Let ConcreteSubjectTest inherit from SubjectTest instead of TestCase 3. Let ConcreteSubjectTest implement makeConcreteSubject() and modify testConcreteSubject()

ภาพที่ 6: รีแฟกทอริงกรณีทดสอบ[5]

3. วิธีการดำเนินงานวิจัย

งานวิจัยนี้มีวัตถุประสงค์ในการเปรียบเทียบความพยายามและความถูกต้องของการสร้างกรณีทดสอบระดับยูนิตสำหรับซอฟต์แวร์ที่ประยุกต์ใช้ดีไซน์แพดเทิร์น ระหว่างวิธีสร้างกรณีทดสอบตามปกติ และวิธีสร้างกรณีทดสอบด้วยการรีแฟกทอริง โดยผู้วิจัยเลือกใช้แผนแบบการทดลองแบบสองกลุ่ม (Randomized Control Group Posttest-Only Design) ซึ่งเป็นแบบแผนการทดลองที่เหมาะสมกับการทดลองที่ต้องการวัดค่าตัวแปรตามความพยายามในการสร้างกรณีทดสอบ และความถูกต้องของกรณีทดสอบจากวิธีสร้างกรณีทดสอบ ของหน่วยการทดลองสองกลุ่มที่ทำการสร้างกรณีทดสอบด้วยวิธีที่แตกต่างกันสองวิธี เพื่อเปรียบเทียบค่าเฉลี่ยความแตกต่างความพยายามและความถูกต้องของวิธีการสร้างกรณีทดสอบ

ความพยายามในงานวิจัยนี้ หมายถึงเวลาของการสร้างกรณีทดสอบหรือรีแฟกทอริงกรณีทดสอบจนถึงการทำความเข้าใจซอฟต์แวร์ เนื่องจากเวลาเป็นปัจจัยอย่างหนึ่งที่สามารถบ่งบอกความพยายามที่ใช้ในการพัฒนาซอฟต์แวร์ได้ [11] [12]

ความถูกต้องของกรณีทดสอบในงานวิจัยนี้หมายถึงร้อยละอัตราส่วนกรณีทดสอบที่สามารถทำงานได้เปรียบเทียบกับจำนวนกรณีทดสอบทั้งหมด ที่ต้องทำการทดสอบ โดยใช้อีซีแอลเอ็มมา (EclEmma) สำหรับเป็นเครื่องมือวัดความครอบคลุมซอร์สโค้ด โดยวัดความครอบคลุมซอร์สโค้ด 3 ด้าน ดังนี้ (1) การครอบคลุมคำสั่งและบรรทัดใน โปรแกรม (2) การครอบคลุมทางแยก (3) การครอบคลุมเงื่อนไข แสดงผล

วิเคราะห์เป็นร้อยละความครอบคลุม และไฮไลท์ซอร์สโค้ด แสดงความครอบคลุมคำสั่งในโปรแกรม การครอบคลุมทางแยก และการครอบคลุมเงื่อนไข

3.1 สมมติฐานการวิจัย

สมมติฐานสองข้อที่งานวิจัยนี้มุ่งทดสอบ คือ (1) ความพยายามจากวิธีสร้างกรณีทดสอบมีความแตกต่างกัน ระหว่างวิธีสร้างกรณีทดสอบตามปกติและวิธีการสร้างกรณีทดสอบด้วยการรีแฟกทอริง (2) ความถูกต้องของกรณีทดสอบมีความแตกต่างกัน ระหว่างวิธีสร้างกรณีทดสอบตามปกติ และวิธีปรับปรุงกรณีทดสอบด้วยการรีแฟกทอริง

3.2 ประชากรและหน่วยทดลอง

การศึกษานี้ผู้วิจัยได้กำหนดประชากร คือ ผู้พัฒนาซอฟต์แวร์ที่มีความรู้ความเข้าใจในการพัฒนาซอฟต์แวร์ด้วยโปรแกรมเชิงวัตถุ มีประสบการณ์ในการเขียนซอฟต์แวร์ด้วยภาษาจาวา และมีความรู้ความเข้าใจดีไซน์แพดเทิร์น ดังนั้นผู้วิจัยจึงกำหนดหน่วยตัวอย่างในการศึกษานี้คือ นิสิตปริญญาโท สาขาวิชาการพัฒนาซอฟต์แวร์ด้านธุรกิจ คณะพาณิชยศาสตร์และการบัญชี และ สาขาวิศวกรรมซอฟต์แวร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย จำนวน 60 คน เนื่องจากไม่ทราบจำนวนของหน่วยประชากรทั้งหมด

3.3 วิธีการดำเนินการทดลอง

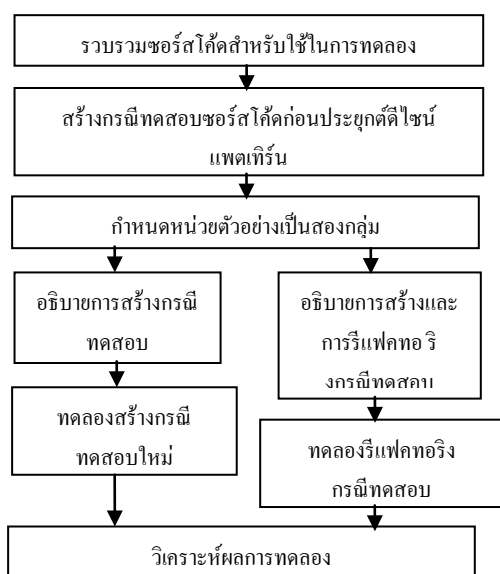
ในการวิจัยนี้จะเลือกศึกษาเฉพาะดีไซน์แพดเทิร์นซึ่งเป็นตัวแทนของดีไซน์แพดเทิร์นแต่ละประเภท ได้แก่ เดคอเรเตอร์ แพดเทิร์น ออบเซิร์ฟเวอร์แพดเทิร์น และแอบสแตรกต์แฟกทอรี แพดเทิร์น เนื่องจากต่างก็เป็นแพดเทิร์นที่มีความซับซ้อน และถูกใช้อย่างแพร่หลาย [5]

ก่อนเริ่มการทดลองนั้น ผู้วิจัยจะแบ่งกลุ่มหน่วยทดลองเป็นสองกลุ่ม กลุ่มละ 30 คน กลุ่มหนึ่งจะได้รับใบงานให้สร้างกรณีทดสอบใหม่จากซอฟต์แวร์ที่ปรับปรุงโดยการรีแฟกทอริงดีไซน์แพดเทิร์น อีกกลุ่มจะได้รับใบงานให้ปรับปรุงกรณีทดสอบด้วยการรีแฟกทอริง

งานวิจัยนี้เลือกใช้เจยูนิทเฟรมเวิร์คสำหรับการทดสอบโปรแกรมภาษาจาวา สำหรับช่วยในการเขียนกรณีทดสอบ และดำเนินการทดสอบ (Run) และเป็นซอฟต์แวร์โอเพนซอร์สที่ถูกใช้อย่างแพร่หลายจนกลายเป็นมาตรฐานในการพัฒนากรณี

ทดสอบและทำการทดสอบยูนิทในภาษาจาวา [5][7] ถูกพัฒนาโดย Erich Gamma และ Kent Beck

ก่อนเริ่มการทดลอง จะมีการอธิบายหลักการของกรณีทดสอบและเฟรมเวิร์กยูนิทสำหรับสร้างกรณีทดสอบให้กับหน่วยทดลองทั้งสองกลุ่ม และอธิบายหลักการรีแฟกทอริงกรณีทดสอบให้เฉพาะกลุ่มที่กำหนดให้รีแฟกทอริงกรณีทดสอบแล้วจึงแจกใบงานให้กับหน่วยทดลองทั้งสองกลุ่ม เพื่อกำหนดให้หน่วยทดลองสร้างกรณีทดสอบคลาสและเมธอดด้วยค่าต่างๆตามที่กำหนดไว้ โดยไม่อธิบายวัตถุประสงค์ที่แท้จริงของการทดลอง



ภาพที่ 7: ภาพแสดงขั้นตอนแนวทางการดำเนินการทดลอง

ในระหว่างการทดลองนั้น ผู้วิจัยจะจับภาพหน้าจอการสร้างกรณีทดสอบไว้เพื่อใช้วิเคราะห์ความพยายามจากการสร้างกรณีทดสอบ และเก็บซอร์สโค้ดเพื่อวิเคราะห์ความถูกต้องจากซอร์สโค้ดกรณีทดสอบ

3.4 ประเด็นความเชื่อถือได้ และความถูกต้อง

เพื่อตอบวัตถุประสงค์ของงานวิจัยนี้ผู้วิจัยได้ควบคุมปัจจัยที่อาจส่งผลต่อการทดลองให้คงที่ภายใต้สภาวะเดียวกัน ดังนี้ (1) ซอฟต์แวร์ที่ต้องทดสอบ หน่วยทดลองทั้งสองกลุ่มทดลองบนซอฟต์แวร์เดียวกันเพื่อป้องกันความยากง่ายที่เกิดจากความแตกต่างกันของซอฟต์แวร์ (2) การสร้างกรณีทดสอบ หน่วยทดลองทั้ง 2 กลุ่มถูกกำหนดให้ทำตามใบงานซึ่งเป็นกรณีทดสอบเดียวกัน มีความครอบคลุมเท่ากัน (3) การวัดความพยายาม เวลาที่ใช้สร้างกรณีทดสอบเป็นตัวแปรสำคัญที่มีผลต่อ

งานวิจัย ซึ่งได้มาจากการจับภาพหน้าจอขณะสร้างกรณีทดสอบของหน่วยทดลอง ผู้วิจัยได้ควบคุมให้ทำการทดลองในห้องปฏิบัติการ (4) การกำหนดหน่วยทดลอง ผู้วิจัยคัดเลือกหน่วยทดลองที่มีพื้นฐานการโปรแกรมด้วยเกณฑ์เดียวกัน

4. ผลการวิเคราะห์ข้อมูล

ข้อมูลที่ได้จากการทดลองแสดงในตารางที่ 1 และ 2 โดยเปรียบเทียบค่าเฉลี่ยของความพยายามในรูปของเวลาที่ใช้ในการสร้างกรณีทดสอบ และความถูกต้องของวิธีการสร้างกรณีทดสอบ

ตารางที่ 1: ตารางแสดงเวลาเฉลี่ยการสร้างกรณีทดสอบทั้งสองวิธี

วิธีสร้างกรณีทดสอบ	เคอเรเตอร์แพดเทิร์น (นาท)	ออปเซิร์ฟเวอร์แพดเทิร์น (นาท)	แอปสเตรกแพดเทิร์น (นาท)
สร้างใหม่	17.7103	21.4143	22.1767
รีแฟกทอริง	9.0263	23.0607	18.7240

ตารางที่ 2: ตารางแสดงความถูกต้องการสร้างกรณีทดสอบทั้งสองวิธี

วิธีสร้างกรณีทดสอบ	เคอเรเตอร์แพดเทิร์น (%)	ออปเซิร์ฟเวอร์แพดเทิร์น (%)	แอปสเตรกแพดเทิร์น (%)
สร้างใหม่	100.00	77.39	89.33
รีแฟกทอริง	100.00	80.00	86.66

ด้วยการทดสอบที (t-test) เพื่อเปรียบเทียบระหว่างหน่วยทดลองกลุ่มที่สร้างกรณีทดสอบใหม่ และหน่วยทดลองที่ใช้วิธีรีแฟกทอริงกรณีทดสอบ ได้ผลการวิเคราะห์ดังนี้

- 1) เวลาที่ใช้ในการสร้างกรณีทดสอบ ระหว่างวิธีสร้างกรณีทดสอบตามปกติและวิธีการสร้างกรณีทดสอบด้วยการรีแฟกทอริง มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05
- 2) ความถูกต้องของกรณีทดสอบที่ได้จากวิธีสร้างกรณีทดสอบตามปกติ และวิธีสร้างกรณีทดสอบด้วยการรีแฟกทอริง ไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05

เมื่อวิเคราะห์แต่ละดีไซน์แพดเทิร์นแยกกัน พบว่าเวลาที่ใช้ในการสร้างกรณีทดสอบในแต่ละดีไซน์แพดเทิร์นต่างๆ

เป็นดังนี้ (1) เดคอเรเตอร์แพตเทิร์น หน่วยทดลองใช้ความพยายามแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ 0.05 (2) ออบเจกต์แพตเทิร์น หน่วยทดลองใช้ความพยายามที่ไม่มีความแตกต่างกัน อย่างมีนัยสำคัญทางสถิติที่ 0.05 (3) แอบสแตรกแพตเทิร์น หน่วยทดลองใช้ความพยายามที่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ 0.05

ส่วนความถูกต้องของกรณีทดสอบในดีไซน์แพตเทิร์นต่างๆ เป็นดังนี้ (1) เดคอเรเตอร์แพตเทิร์น หน่วยทดลองสามารถสร้างกรณีทดสอบได้ถูกต้องไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ 0.05 (2) ออบเจกต์แพตเทิร์น หน่วยทดลองสร้างกรณีทดสอบได้ถูกต้องแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ 0.05 (3) แอบสแตรกแพตเทิร์น หน่วยทดลองสร้างกรณีทดสอบที่มีความถูกต้องแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ 0.05

5. สรุปและอภิปรายผล

5.1 ความพยายามในการสร้างกรณีทดสอบ

ผลการวิเคราะห์ข้อมูล พบว่า ความพยายามในการสร้างกรณีทดสอบ ระหว่างวิธีสร้างกรณีทดสอบตามปกติและวิธีการสร้างกรณีทดสอบด้วยการรีแฟกทอริง มีความแตกต่างกัน โดยเฉพาะอย่างยิ่ง การรีแฟกทอริงกรณีทดสอบสำหรับเดคอเรเตอร์แพตเทิร์น และแอบสแตรกแพตเทิร์น ดังนั้น ความพยายามในการสร้างกรณีทดสอบขึ้นอยู่กับดีไซน์แพตเทิร์นที่ใช้ด้วย บางแพตเทิร์นส่งผลใช้ความพยายามมากขึ้น แต่บางแพตเทิร์นให้ส่งผลใช้ความพยายามลดลง เนื่องจากการประยุกต์บางแพตเทิร์นส่งผลต่อจำนวนคลาส และเมธอดที่ต้องทดสอบเพิ่มขึ้น จึงทำให้ต้องใช้พยายามรีแฟกทอริงกรณีทดสอบเพิ่มขึ้นด้วย

5.2 ความถูกต้องของกรณีทดสอบ

ผลการวิเคราะห์ข้อมูล พบว่า ไม่มีความแตกต่างของความถูกต้องของกรณีทดสอบ ระหว่างวิธีสร้างกรณีทดสอบตามปกติและวิธีการสร้างกรณีทดสอบด้วยการรีแฟกทอริง แต่ในขณะที่เดียวกันพบว่าความถูกต้องของกรณีทดสอบต่อออบเจกต์แพตเทิร์น การรีแฟกทอริงให้ความถูกต้องสูงกว่า เนื่องจากการประยุกต์ใช้ดีไซน์แพตเทิร์นมีผลลดความซับซ้อนของกรณีทดสอบ ส่วนความถูกต้องของกรณีทดสอบต่อแอบส

แตรกแพตเทิร์นแพตเทิร์นนั้น การรีแฟกทอริงให้ความถูกต้องน้อยกว่า เนื่องจาก การประยุกต์ใช้ดีไซน์แพตเทิร์น มีโครงสร้างมีโครงสร้างที่ต้องทดสอบซับซ้อนมากขึ้น

5.3 การนำงานวิจัยไปประยุกต์ใช้และโอกาสการศึกษาในอนาคต

การค้นพบในงานวิจัยชิ้นนี้ ในเชิงทฤษฎีสามารถช่วยยืนยันเชิงประจักษ์ให้ชัดเจนมากขึ้นถึงวิธีสร้างกรณีทดสอบที่มีผลต่อความพยายาม แต่ไม่มีผลต่อความถูกต้อง และในเชิงประยุกต์สามารถช่วยให้ผู้ทดสอบซอฟต์แวร์สามารถเลือกใช้วิธีสร้างกรณีทดสอบที่เหมาะสมได้ เนื่องจากงานวิจัยนี้ทำการทดลองเปรียบเทียบการสร้างกรณีทดสอบจากเพียงสามแพตเทิร์นเท่านั้น ดังนั้นข้อเสนอแนะสำหรับงานวิจัยในอนาคต อาจศึกษาแพตเทิร์นอื่นๆ หรือพิจารณาตัวแปรอื่นเพิ่มเติม เช่นขนาดของซอฟต์แวร์ที่ใหญ่ขึ้น และมีความซับซ้อนมากขึ้น เพื่อทราบถึงผลลัพธ์ในบริบทอื่นๆ ที่สะท้อนความเป็นจริงมากขึ้น

เอกสารอ้างอิง

- [1] Robert C. Martin, *Agile Software Development*, Prentice Hall: United State, 2002.
- [2] Joshua Kerievsky, *Refactoring to Patterns*, Pearson Education: United State, 2005.
- [3] G.B. Regulwar and R.M. Tagnayat, "Bad Smelling Concept in Software Refactoring," *International Proceedings of Economics Development & Research*, Vol. 45, pp. 56, August 2012.
- [4] M. Fowler, B. Kent, *Refactoring Improving the Design of Existing Code*, Addison-Wesley: United State, 1999.
- [5] P.-H. Chu, N.-L. Hsueh, H.-H. Chen and C.-H. Liu, "A Test Case Refactoring Approach for Pattern-Based Software Development" *Springer Trans. on Software Qual J*, vol. 20, pp. 43-75, March 2012.
- [6] E. Gamma, R. Helm, R. Johnson and J. Vlissides, *Design Patterns: Element of Reusable Object-Oriented Software*, Addison-Wesley: United State, 1994.
- [7] P. Tahchiev, F. Leme, V. Massol and G. Gregory, *JUnit in Action*, Manning Publication: United State, 2010.
- [8] D. Jungol, D. Buchmann and U. Ultes-Nitsche, "Assessment of Code Quality through Classification of Unit Tests in VeriNec," *AINAW '07 Proceedings of the 21st International Conference on Advanced Information Networking and Application Workshop*, Vol. 01, pp. 177-182, 2007.

- [9] H. Zhu, P. Hall and J. May, "Software Unit Test Coverage and Adequacy" *ACM Computing Surveys*, Vol. 29, pp. 366-427, 1997.
- [10] I. Khriiss, R.K. Keller and I.A. Hamid, "Pattern-based Refinement Schemas for Design Knowledge Transfer," *Knowledge-Based Systems*, Vol. 13, pp. 403-415, 2000.
- [11] R. S. Pressman. *Software Engineering: A Practitioner's Approach*, McGraw-Hill: New York, 2004.
- [12] L. Hochstein, V.R. Basili, U. Vishkin, J. Gilbert , "A Pilot Study to Compare Programming Effort for Two Parallel Programming Models," *Journal of Systems and Software*, Vol. 81, Issue 11, pp. 1920-1930, November 2008.

การระบุตัวแทนสินทรัพย์ทั่วไปในซอร์สโค้ด ด้วยการเปรียบเทียบเส้นทางของการทดสอบซอฟต์แวร์ Identifying Common Asset Candidates in Source Code by Comparing Software Testing Paths

สิทธิพล ลิ้มชัยชะดา (Sitthipon Limchaichada)¹ และ พรศิริ หมั่นไชยศรี (Pornsiri Muenchaisri)²

ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

¹Sitthipon.L@student.chula.ac.th, ²Pornsiri.Mu@chula.ac.th

บทคัดย่อ

การระบุตัวแทนสินทรัพย์ทั่วไปเป็นการนำเอาส่วนประกอบร่วมของผลิตภัณฑ์ในอดีตที่มีความสามารถในการตอบสนองต่อความต้องการในขอบเขตของการทำงานเดียวกันมาแปลงให้เป็นสินทรัพย์หลักโดยอยู่บนพื้นฐานของการนำซอฟต์แวร์มาใช้ซ้ำ ในปัจจุบันนั้นการระบุตัวแทนสินทรัพย์ทั่วไปในประเภทของซอร์สโค้ดที่มีการนิยามเชิงพฤติกรรมที่เหมือนกันแต่วิธีการเขียนที่แตกต่างกันนั้นยังมีข้อจำกัด เนื่องจากเทคนิคการตรวจหาสำเนาโค้ดโดยส่วนใหญ่อาศัยความสัมพันธ์ของโครงสร้าง ในการตรวจหา ซึ่งไม่สามารถตรวจหาสำเนาโค้ดในนิยามเชิงพฤติกรรมได้ ดังนั้นงานวิจัยนี้มีจุดประสงค์เพื่อการแก้ปัญหาในการระบุตัวเลือกสินทรัพย์ร่วมในระดับโค้ดดังกล่าวโดยวิธีการตรวจหาความเหมือนของโค้ดที่ใช้วิธีการเขียนแตกต่างกัน แต่ให้พฤติกรรมหรือผลลัพธ์การทำงานที่เหมือนกัน (สำเนาโค้ดประเภทที่ 4) โดยใช้เส้นทางของวิธีการทดสอบซอฟต์แวร์ จากผลการทดลองพบได้ว่าการระบุตัวเลือกสินทรัพย์ร่วมในระดับโค้ดด้วยวิธีดังกล่าวสามารถสร้างตัวแทนสินทรัพย์ทั่วไปเพื่อนำมาเป็นสินทรัพย์หลักในการพัฒนาผลิตภัณฑ์ที่มีลักษณะการทำงานที่เหมือนกันของเมทอดที่จะเกิดขึ้นในอนาคต

คำสำคัญ: การพัฒนาสินทรัพย์ทั่วไป สินทรัพย์ทั่วไป กราฟควบคุมสายงาน การทดสอบเส้นทาง

Abstract

Identifying common assets is bringing common components of the product in the past with the ability to

meet requirements in the scope of work to convert it to core asset base on the basis of introducing software reuse. The current identifying common assets of the types of source code fragments that perform the same function but how to implemented code different have limitations because the techniques to detect duplicate code by relying mostly associated structures to detect which can not detect a behavioral code definitions. The purpose of this research solve the problem of identifying a common asset of source code fragments that perform the same function but how to implemented code different (Type 4) by method of comparing software testing paths. The result shows that specifying common assets in the level code can create common asset to be core asset in the development of products with the same behavior of the methods that will occur in the future.

Keywords: Core Asset Development, Common Asset ,Control Flow graph ,Path Testing

1. คำนำ

การพัฒนาสินทรัพย์หลัก(Core asset development)[1] เป็นหนึ่งในกิจกรรมที่มีความสำคัญมากในการสนับสนุนให้เกิดกระบวนการในการผลิตสินค้าให้มีความหลากหลายและปรับให้เข้ากับความต้องการของผู้บริโภคที่มีได้เป็นอย่างดี เนื่องมาจากการพัฒนาสินทรัพย์หลักนั้นเป็นการนำเอาส่วนประกอบร่วม(Common component)ของผลิตภัณฑ์ในอดีตที่มีความสามารถในการตอบสนองต่อความต้องการใน

ขอบเขตของการทำงานเดียวกันมาแปลงให้เป็นสินทรัพย์หลัก (core asset base) โดยอยู่บนพื้นฐานของการนำซอฟต์แวร์มาใช้ซ้ำ (Software reuse) และตัวอย่างของการพัฒนาซอฟต์แวร์ที่มีการสนับสนุนการนำกลับมาใช้ซ้ำอย่างเป็นระบบในปัจจุบัน เพื่อตอบสนองความหลากหลายของผลิตภัณฑ์คือแนวคิดของสายผลิตภัณฑ์ซอฟต์แวร์ (Software product lines) ที่มีการพัฒนาสินทรัพย์หลัก (core asset) สำหรับนำไปใช้ซ้ำในการประกอบเป็นผลิตภัณฑ์

การตรวจหาสำเนาโค้ดเป็นหนึ่งในเทคนิคที่มีความเหมาะสมต่อการระบุส่วนร่วมของผลิตภัณฑ์ การตรวจหาสำเนาโค้ด คือการตรวจหาชิ้นส่วนโค้ดที่เหมือนหรือคล้ายกัน หรือเรียกว่า สำเนาโค้ด (code clone) ซึ่งมักพบได้ทั่วไปในการพัฒนาซอฟต์แวร์ [3], [4] ทำให้การตรวจหาสำเนาโค้ดถูกนำไปใช้ในหลายขอบเขตและวัตถุประสงค์ แม้ว่าสำเนาโค้ดมักจะเกิดจากการทำสำเนาและวาง (copy and paste) แต่นักเขียนโปรแกรมส่วนใหญ่จะทำการดัดแปลงสำเนาโค้ดเพื่อนำไปใช้ซ้ำทำให้เกิดความแตกต่างระหว่างคู่สำเนาโค้ด ซึ่งเป็นอุปสรรคต่อการสร้างสินทรัพย์ทั่วไปจากสำเนาโค้ดที่ตรวจหาได้ว่าจะตัดสินใจแยกความแตกต่างเหล่านั้นออกมาอย่างไร

ในบทความ [6] ได้มีวิธีการออกแบบวิธีการและพัฒนาเครื่องมือระบุตัวเลือกสินทรัพย์ร่วมจากซอร์สโค้ดระหว่างผลิตภัณฑ์โดยประยุกต์ใช้เทคนิคการตรวจหาสำเนาโค้ดในการระบุส่วนร่วมและใช้แบบอย่างการออกแบบประเภทแม่ที่ออกแบบในการแยกความแตกต่างออกจากสำเนาโค้ดที่ตรวจพบและก่อนเป็นตัวเลือกสินทรัพย์ร่วม อย่างไรก็ตามบทความ [6] เสนอเพียงสำเนาโค้ดประเภทที่ 3 (รูปแบบของสำเนาโค้ดได้กล่าวไว้ในหัวข้อ 3.5) เท่านั้นว่าจะทำอย่างไรกับโค้ดที่ทำการเปลี่ยนแปลงและเพิ่มลบคำสั่ง แต่ไม่ได้รวมถึงสำเนาโค้ดประเภทที่ 4 ใช้วิธีการเขียนแตกต่างกันแต่ให้พฤติกรรมหรือผลลัพธ์การทำงานที่เหมือนกัน

งานวิจัยนี้ได้แนะนำการแก้ปัญหาในการระบุตัวเลือกสินทรัพย์ร่วมในระดับโค้ดดังกล่าวโดยวิธีการตรวจหาความเหมือนของโค้ดที่ใช้วิธีการเขียนแตกต่างกัน แต่ให้พฤติกรรมหรือผลลัพธ์การทำงานที่เหมือนกัน (สำเนาโค้ดประเภทที่ 4) โดยใช้วิธีการเปรียบเทียบเส้นทางของการทดสอบซอฟต์แวร์เข้ามาประยุกต์ใช้ในขั้นตอนของกระบวนการที่ออกแบบโดย

การตรวจสอบความเหมือนของโครงสร้างและพฤติกรรมก่อนที่จะนำไปก่อตัวแบบตัวเลือกสินทรัพย์

สำหรับรายละเอียดในส่วนต่อไปในหัวข้อที่ 2 และ 3 จะกล่าวถึงงานวิจัยและทฤษฎีที่เกี่ยวข้องตามลำดับ ในหัวข้อที่ 4 จะกล่าวถึง วิธีการดำเนินงานวิจัย หัวข้อที่ 5 กล่าวถึงการทดลองและการประเมินผล และหัวข้อสุดท้ายคือหัวข้อที่ 6 เป็นการสรุปผลงานวิจัยและแนวทางดำเนินการในอนาคต

2. งานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ทำการศึกษางานวิจัยที่เกี่ยวข้องกับการระบุตัวแทนสินทรัพย์ทั่วไปในซอร์สโค้ด และคุณสมบัติของ Matching Technique โดยมีวัตถุประสงค์เพื่อเป็นแนวทางในการทำงานวิจัยให้มีประสิทธิภาพให้เกิดผลลัพธ์ที่มีความสมบูรณ์

บทความ [6] ได้มีการนำเสนอแนวคิดที่ถูกทำขึ้นเพื่อลดค่าใช้จ่ายและแรงงานในการพัฒนาซอฟต์แวร์ซึ่งได้มีการพัฒนาหลักสินทรัพย์เพื่อนำไปใช้ในการประกอบและปรับแต่งเป็นส่วนหนึ่งของผลิตภัณฑ์ที่จะทำการสร้างขึ้นมา ซึ่งในบทความนี้ได้มีการประยุกต์ใช้การตรวจหาสำเนาโค้ด ระหว่างผลิตภัณฑ์เพื่อระบุตัวแทนสินทรัพย์ทั่วไปจากผลิตภัณฑ์ที่มีอยู่ โดยใช้การตรวจหาสำเนาโค้ดประเภทที่ 3 คือ กรณีที่สำเนาโค้ดที่ตรวจพบมีจุดแตกต่างอยู่บ้าง จุดแตกต่างนั้นจำเป็นต้องถูกแยกออกมาเพื่อให้สามารถนำสำเนาโค้ดนั้นไปพัฒนาเป็นสินทรัพย์ทั่วไปได้อย่างไรก็ตามบทความ [6] เสนอเพียงสำเนาโค้ดประเภทที่ 3 เท่านั้นว่าจะทำอย่างไรกับโค้ดที่ทำการเปลี่ยนแปลง และเพิ่มลบคำสั่ง แต่ไม่ได้รวมถึงสำเนาโค้ดประเภทที่ 4 ที่ชิ้นส่วนโค้ดที่ใช้วิธีการเขียนแตกต่างกันแต่ให้พฤติกรรมหรือผลลัพธ์การทำงานที่เหมือนกัน

บทความงานวิจัยของ Miryung Kim, David Notkin [8] ได้รวบรวมและสำรวจเกี่ยวกับความสามารถและคุณสมบัติของ Matching Technique ซึ่งเป็นเทคนิคที่ใช้ในการวิเคราะห์เวอร์ชันของซอฟต์แวร์โดยเทคนิคต่างๆ เช่น Entity Name Matching, String Matching, Syntax Tree Matching, Control Flow Graph Matching เป็นต้น การทดลองและสรุปผลได้ว่าในแต่ละวิธีนั้นจะดีและข้อเสียแตกต่างกันไปขึ้นอยู่กับองค์ประกอบต่างๆ ของโปรแกรมที่มีอยู่

นอกจากนี้ NiCad [12] เป็นเครื่องมือตรวจหาสำเนาโค้ดซึ่งใช้เทคนิคการเปรียบเทียบเชิงข้อความทำให้ใช้ทรัพยากรในการตรวจหาที่ค่อนข้างต่ำเมื่อเทียบกับเทคนิคประเภทอื่นโดยใช้

ภาษาที่เอ็กซ์แอล(TXL) เป็นกลไกสำคัญในการแปลงโค้ดเพื่อใช้เปรียบเทียบชิ้นส่วน

3. ทฤษฎีที่เกี่ยวข้อง

3.1 การพัฒนาสินทรัพย์ (Core asset development)

[1]

การสร้างสินทรัพย์(core asset base)ซึ่งเป็นการนำส่วนประกอบทั่วไป(common component)มาแปลงให้เป็นสินทรัพย์(core asset base) โดยอยู่บนพื้นฐานของการนำกลับมาใช้ใหม่(reuse) แต่การนำกลับมาใช้ใหม่(reuse) ต้องเป็นการนำกลับมาใช้ใหม่(reuse)

3.2 กราฟการไหลของการควบคุม (Control Flow graphs) [5]

กราฟควบคุมสายงาน(Control Flow graphs) เป็นการแสดงโครงสร้างการควบคุมของโปรแกรมในลักษณะเชิงโครงสร้างของโปรแกรมซึ่งคล้ายกับแผนภาพโฟลว์ชาร์ท แต่สามารถไม่ต้องใส่รายละเอียดทุกคำสั่งของขบวนการ

3.3 การทดสอบเส้นทาง (Path Testing) [5]

การทดสอบเส้นทาง(Path Testing) เป็นวิธีการหนึ่งสำหรับการทำการทดสอบขนาดย่อย(Unit Testing) ซึ่งอยู่ในรูปแบบประเภทการ ทดสอบ แบบกล่องขาว (White Box) เพื่อต้องการตรวจการดำเนินงานของโปรแกรมว่าครอบคลุมทุกคำสั่งการทำงานหรือไม่(Coverage Testing) ดังนั้นแล้วการทดสอบเส้นทาง(Path Testing) จึงสามารถตรวจสอบการเขียนโปรแกรมของโปรแกรมเมอร์ หรือนักพัฒนา(Developer) ตรงกับความต้องการของผู้ใช้ที่ได้กำหนดสเปคของโปรแกรม(Program Specification) หรือไม่

3.4 การตรวจหาสำเนาโค้ด [2]

สำเนาโค้ด คือ ชิ้นส่วนซอร์สโค้ดที่ถูกสำเนาเพื่อการใช้งาน และอาจทำการแก้ไขให้เข้ากับการใช้งานใหม่ รายละเอียดของสำเนาโค้ดทั้ง 4 ประเภท ได้แก่

- 1).ประเภทที่ 1 เป็นชิ้นส่วนโค้ดที่ถูกคัดลอกมาโดยสมบูรณ์ไม่มีความแตกต่างนอกจากช่องว่าง(whitespace) และหมายเหตุ (comment)
- 2).ประเภทที่ 2 ชิ้นส่วนโค้ดยังมีความคล้ายกันในเชิงวากยสัมพันธ์หรือเชิงโครงสร้าง แต่อาจมีการเปลี่ยนแปลงชื่อในส่วนตัวระบุ(identifier) สัญญาพจน์(literal) หรือแบบชนิด(type) รวมถึงช่องว่าง และหมายเหตุ

3).ประเภทที่ 3 ชิ้นส่วนโค้ดมีการแก้ไข โดยการเพิ่ม ลบ หรือเปลี่ยนแปลงข้อความสั่ง(statement) ทำให้เกิดความแตกต่างในบรรทัดใดๆ ในชิ้นส่วนโค้ด สำเนาโค้ดชนิดนี้อาจกล่าวได้ว่าเป็นสำเนาประเภทที่ 1 ซึ่งมีความแตกต่างแทรกอยู่ในชิ้นส่วนโค้ด

4).ประเภทที่ 4 ชิ้นส่วนโค้ดใช้วิธีการเขียนที่แตกต่างกัน แต่ให้พฤติกรรมหรือผลลัพธ์การทำงานที่เหมือนกัน เช่น for<>recursive , if-else<>switch case , while<>recursive เป็นต้น

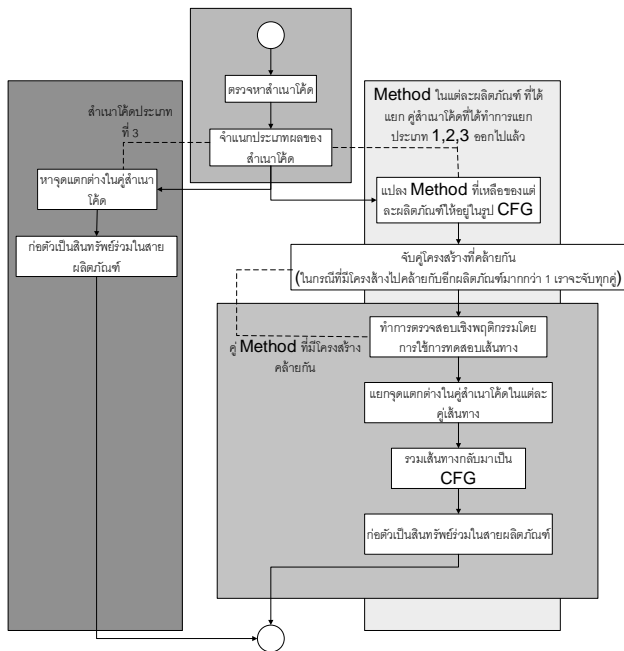
ภาพที่ 1: ตัวอย่างสำเนาโค้ดเชิงข้อความ ประเภทที่ 1,2,3,4

3.5 เมทอดแม่แบบ[6]

เมทอดแม่แบบเป็นรูปแบบการออกแบบอย่างหนึ่งซึ่งใช้ในการออกแบบเชิงวัตถุ โดยมีจุดประสงค์เพื่อลดความซ้ำซ้อนในกรณีที่มีเมทอดนั้นๆ มีทั้งโค้ดส่วนที่เหมือนกันและแตกต่างกัน จึงใช้เมทอดแม่แบบเพื่อแยกส่วนทั้งสองที่แตกต่างกันออกไป ทำให้สามารถนำโค้ดอื่นๆ ที่เหมือนกันไปรวมไว้ยังคลาสแม่ได้

4. ขั้นตอนของขบวนการระบุตัวเลือกสินทรัพย์ร่วมของโค้ดที่ใช้วิธีการเขียนแตกต่างกัน แต่ให้พฤติกรรมหรือผลลัพธ์การทำงานที่เหมือนกัน

การออกแบบขบวนการของการแก้ไขปัญหาแบ่งออกเป็น 2 ขั้นตอนคือ การตรวจหาสำเนาโค้ดที่มีลักษณะของโครงสร้างการทำงานที่เหมือนกันโดยการเปรียบเทียบกราฟการไหลของการควบคุม และ การระบุเป็นสำเนาโค้ดประเภทที่ 4 เพื่อออกเป็นสินทรัพย์ร่วมในผลิตภัณฑ์ ดังแสดงในภาพรวมของงานวิจัยภาพที่ 2



ภาพที่ 2: ภาพรวมเปรียบเทียบขั้นตอนของงานวิจัยในปัจจุบันและงานวิจัยที่นำมาใช้รวมในอดีต

4.1 การออกแบบขั้นตอนของการตรวจสอบสำเนาโค้ดที่มีลักษณะของโครงสร้างการทำงานที่เหมือนกันโดยการเปรียบเทียบกราฟการไหลของการควบคุม

ในขั้นตอนนี้จะแบ่งออกเป็นทั้งหมด 3 ขั้นตอนเพื่อที่จะนำเมทอดที่เหลือจากการคัดแยกประเภทที่ 1,2,3 ออกเพื่อที่จะเข้าสู่ขบวนการเพื่อที่จะออกมาเป็นคู่เมทอดที่มีโครงสร้างคล้ายกัน

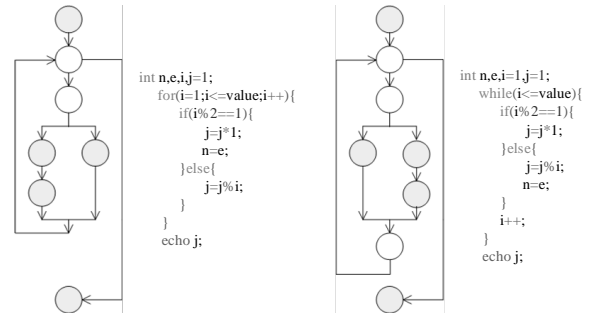
1).การแปลงโค้ดที่เหลืออยู่ทุกเมทอดของทั้งสองผลิตภัณฑ์ให้อยู่ในกราฟการไหลของการควบคุมแล้วทำการปรับกราฟให้อยู่ในรูป DD-Path (Decision-to-Decision path) เพื่อจัดโครงสร้างให้สามารถเข้าใจได้ง่ายเพื่อรวมคำสั่งที่เป็นลักษณะเส้นตรงมารวมเป็นจุดเดียวกันเพื่อง่ายต่อการเปรียบเทียบ

2).การเปรียบเทียบความเหมือนของกราฟการไหลของการควบคุม เพื่อให้แน่ใจว่าทั้งสองเมทอดมีโครงสร้างที่เหมือนกันโดยการเปรียบเทียบจุดการทำงานเป็นจุดต่อจุดว่า มีเส้นทางเข้าและออกในจำนวนที่เท่ากันหรือไม่ ในส่วนนี้จะมีการเปรียบเทียบโครงสร้างโดยรวมที่ง่ายขึ้นเนื่องจากกราฟการไหลของการควบคุมได้มีการแปลงรูปร่างเป็นลักษณะของ DD-Path Decision-to-Decision path

3).การเปรียบเทียบโค้ดในส่วนของจุดตัดสินใจบนกราฟการไหลของการควบคุมเพื่อตรวจสอบโครงสร้างในรูปที่เกิดขึ้นว่ามีพฤติกรรมของการตัดสินใจที่เหมือนกันหรือไม่เช่น

จำนวนรูป ค่าตัวแปรที่จะส่งผลกระทบต่อภายในรูป ค่าเริ่มต้น ค่าของการวนลูป และค่าของการหยุดการวนลูป

ซึ่งเมื่อทำครบทั้ง 3 ขั้นตอนแล้ว จะได้ข้อมูลนำออกเป็นเมทอดที่มีโครงสร้างของการทำงานที่เหมือนกัน(โดยจะยังไม่สนใจต่อพฤติกรรมของการทำงานในขั้นตอนนี้)จะนำไปเป็นข้อมูลนำเข้าในขั้นตอนถัดไป ดังภาพที่ 3



ภาพที่ 3: ผลลัพธ์ของการวิเคราะห์โครงสร้าง

4.2 การออกแบบขั้นตอนของการระบุเป็นสำเนาโค้ดประเภทที่ 4 เพื่อก่อเป็นสินทรัพย์ร่วมในผลิตภัณฑ์

ในส่วนนี้เมื่อทำการตรวจสอบสำเนาโค้ดทั้งสี่ประเภทแล้วจะนำคู่สำเนาโค้ดของทั้ง 2 ผลิตภัณฑ์ที่มีโครงสร้างเหมือนกันมาเป็นข้อมูลนำเข้าแล้วผ่านขบวนการทั้งหมด 4 ขั้นตอน(ภาพที่ 4)ได้แก่

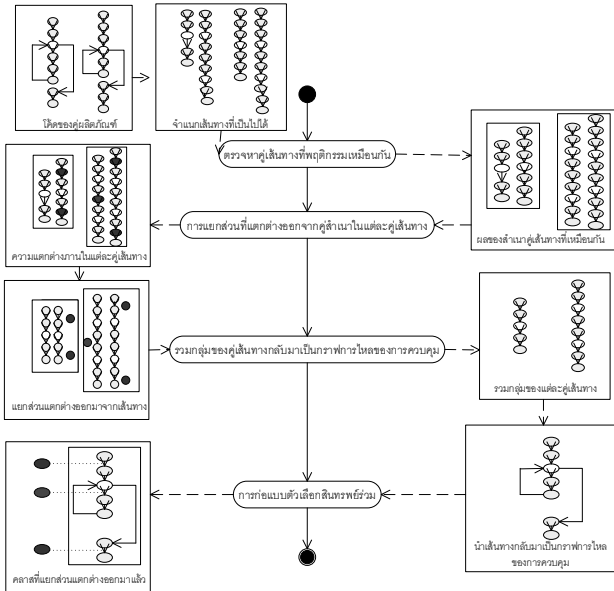
1).การตรวจหาคู่เส้นทางที่ทำงานเหมือนกัน โดยในขั้นตอนนี้จะมีข้อมูลที่เป็นข้อมูลนำเข้า คือ กราฟการไหลของการควบคุม ที่เหมือนกันของทั้งสองเมทอด แล้วทำการแจกแจงเส้นทางที่เป็นไปได้ทั้งสองผลิตภัณฑ์ ต่อมาจะทำการจับคู่เส้นทางที่มีพฤติกรรมของการทำงานที่เหมือนกัน โดยข้อมูลนำออกในขั้นตอนนี้จะเป็น คู่เส้นทางที่มีพฤติกรรมของการทำงานที่เหมือนกัน

2).การแยกส่วนที่แตกต่างออกจากคู่สำเนาในแต่ละคู่เส้นทาง เมื่อได้คู่เส้นทางที่มีพฤติกรรมของการทำงานที่เหมือนกันแล้วต่อมานำเอาแต่ละคู่มาแยกจุดแตกต่างออกจากรวมกัน และจะได้ข้อมูลนำออกเป็นตัวแทนของแต่ละคู่ที่มีการสกัดจุดแตกต่างออกไป

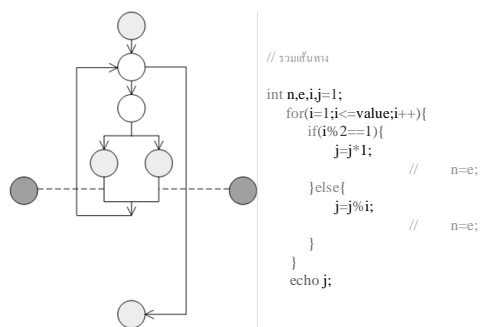
3).การรวมกลุ่มของคู่เส้นทางกลับมาเป็นกราฟการไหลของการควบคุม ในขั้นตอนนี้จะทำการนำตัวแทนของแต่ละคู่ที่มีการสกัดจุดแตกต่างออกไปมารวมกลับเป็นกราฟการไหลของการควบคุม ซึ่งเป็นข้อมูลนำออกในขั้นตอนนี้

4).การก่อแบบตัวเลือกสินทรัพย์ร่วม จะนำเอากราฟการไหลของการควบคุมในขั้นตอนที่ผ่านมาแล้วทำการโอเวอร์ไรด์จุดที่แตกต่างที่สกัดออกมากลับเข้าไป

จนได้ผลลัพธ์ของข้อมูลนำออกจะเป็นสินทรัพย์ร่วมของทั้งสองผลิตภัณฑ์ซึ่งอยู่ในรูปแบบของแม่แบบ ดังภาพที่ 5



ภาพที่ 4: ภาพรวมขั้นตอนการแก้ปัญหา



ภาพที่ 5: ผลลัพธ์ของการก่อสินทรัพย์ร่วมของ 2 แม่ทอด

5. การทดลองและการประเมินผล

การสร้างสินทรัพย์ทั่วไปในผลิตภัณฑ์จะได้ผลลัพธ์สูงสุดก็ต่อเมื่อแม่ทอดทั้งสองที่มีโครงสร้างของการทำงานที่คล้ายกันในเชิงพฤติกรรมสามารถสกัดสิ่งที่เป็นแม่แบบของการทำงานร่วมกันแยกส่วนที่แตกต่างออกจากส่วนทั่วไปนั้นโดยใช้แม่ทอดแม่แบบ

ในงานวิจัยนี้ได้จำลองสถานการณ์ผลิตภัณฑ์จัดการสต็อกสินค้าทั้งสองประเภทที่มีการแทรกโค้ดในลักษณะแม่ทอดของการเขียนโค้ดที่มีพฤติกรรมเหมือนกันแต่เขียนต่างกันไว้ทั้งสองผลิตภัณฑ์เพื่อที่จะทำการทดสอบการจับคู่และสร้างสินทรัพย์ร่วมของทั้งสองผลิตภัณฑ์

จากการทดลองในส่วนแรกพบว่า ในขบวนการเปรียบเทียบความเหมือนของกราฟการไหลของการควบคุมเพื่อให้แน่ใจว่าทั้งสองแม่ทอดมีโครงสร้างที่เหมือนกันได้มีการปรากฏผังตารางที่ 1 ซึ่งในตารางดังกล่าวจะปรากฏให้เห็นว่าจะได้คู่แม่ทอดที่มีลักษณะของโครงสร้างสำเนาโค้ดประเภทที่ 4 ทั้งหมด 12 คู่ ซึ่งแบ่งเป็นในรูปแบบของ If else/switch 6 คู่ For/While 4 คู่ while/recursive 2 คู่

ตารางที่ 1: แสดงผลการเปรียบเทียบการทำงานที่เหมือนกันเชิงโครงสร้าง

ผลิตภัณฑ์	สำเนาโค้ด			สำเนาโค้ดประเภทอื่นๆ ที่มีการทำงานของโครงสร้างที่เหมือนกัน			
	ประเภทที่ 1(คู่)	ประเภทที่ 2(คู่)	ประเภทที่ 3(คู่)	If/else/switch	For/While	while/recursive	Other
ShopA - MartP	7	5	4	6	4	-	2

ตารางที่ 2: แสดงผลการนำกรณีทดสอบของแม่ทอดเดิมของทั้งสองผลิตภัณฑ์ไปประมวลผลกับสินทรัพย์ร่วม

เปรียบเทียบผลิตภัณฑ์	รูปแบบสำเนาโค้ดประเภทที่ 4	ชื่อ method	Te st cas e	Te st cas e	ความถูกต้องของสินทรัพย์ร่วม2 ผลิตภัณฑ์/ test case A		ความถูกต้องของสินทรัพย์ร่วม2 ผลิตภัณฑ์/ test case B	
					ผลิ ตภ ณฑ์	ผลิ ตภ ณฑ์	ผลิ ตภ ณฑ์	ผลิ ตภ ณฑ์
Sh op A- Ma rtP	If else/ switc h	CatA – Swit	6	6	6/6	1	6/6	1
		Stock – Fill	6	6	6/6	1	6/6	1
		Sum – Ship	2	2	2/2	1	2/2	1
		Shelf – Shop	5	5	5/5	1	5/5	1
	Sp1 – po1	7	3	ไม่สามารถสร้างสินทรัพย์ร่วมได้				
For/ whil e		SSop – Kart1	6	2	ไม่สามารถสร้างสินทรัพย์ร่วมได้			
		Sed – Oma	6	6	6/6	1	6/6	1
		Wet1 – Pmar	2	2	2/2	1	2/2	1
		Wet2 – Cpa1	7	7	7/7	1	7/7	1

	CartT – Man	4	2	ไม่สามารถสร้าง สินทรัพย์ร่วมได้			
Whil e/ rec ur	Shop1 – Ma1	7	7	7/7	1	7/7	1
	Shop2 – MT2	9	9	9/9	1	9/9	1

หลังจากนั้นจะเข้าสู่ขบวนการระบุเป็นสำเนาโค้ดประเภทที่ 4 เพื่อก่อเป็นสินทรัพย์ร่วมในผลิตภัณฑ์เมื่อทำสร้างสินทรัพย์ร่วมของทั้ง 3 ผลิตภัณฑ์จะพบว่าบางคู่สำเนาโค้ดที่ไม่สามารถมาสร้างเป็นสินทรัพย์ร่วมได้เช่น If else / switch 2 คู่ For/While 1 คู่ เนื่องจากมีจำนวนลูปที่เท่ากันแต่คำสั่งภายในลูปแตกต่างกัน จึงส่งผลให้มีเส้นทางที่เกิดขึ้นไม่เท่ากันและคำสั่งภายในเส้นทางมีความแตกต่างกันอย่างชัดเจน จึงต้องทำการสกัดออกไป

เมื่อนำไปประเมินผลระหว่างเมทอดเดิมของทั้งสองผลิตภัณฑ์เปรียบเทียบกับสินทรัพย์ร่วมที่เป็นเมทอดแม่แบบโดยการนำกรณีทดสอบของเมทอดเดิมของทั้งสองผลิตภัณฑ์ไปประมวลผลกับสินทรัพย์ร่วม ดังตารางที่ 2 พบว่าผลลัพธ์ของการนำกรณีทดสอบเบื้องต้นของทั้งสองผลิตภัณฑ์ไปทดสอบกับสินทรัพย์ร่วมของทั้งสองผลิตภัณฑ์ ผลลัพธ์ที่ได้จะมีค่าความถูกต้องเฉลี่ยเท่ากับ 1 ซึ่งไม่มีความผิดพลาดที่เกิดขึ้นเลย แต่ในส่วนของกลุ่มเมทอดที่มีความซับซ้อนซึ่งมีความแตกต่างกันอยู่ในส่วนของความแตกต่างที่ไม่มีความเหมือนกันตลอดคู่เส้นทางใดเส้นทางหนึ่งจะถูกสกัดออกในส่วนขั้นตอนที่ 2 จะทำให้ไม่สามารถนำมาสร้างเป็นสินทรัพย์ร่วมได้

6. สรุปผลงานวิจัยและแนวทางการดำเนินการในอนาคต

งานวิจัยนี้สามารถนำเสนอแนวทางจัดการกับชิ้นส่วนโค้ดที่มีวิธีการเขียนที่แตกต่างกันแต่ให้พฤติกรรมที่เหมือนกันที่ตรวจหาได้โดยใช้วิธีการของการทดสอบซอฟต์แวร์เพื่อใช้สร้างเป็นตัวแทนสินทรัพย์ทั่วไปทำให้ลดภาระของผู้เชี่ยวชาญในการวิเคราะห์โค้ดของผลิตภัณฑ์เพื่อระบุตัวแทนสินทรัพย์ทั่วไป

ผลการประเมินของงานวิจัยนี้พบว่า การทดสอบซอฟต์แวร์สามารถตรวจสอบพฤติกรรมของการทำงานของโค้ดในระดับของเมทอดโค้ดในทุกกรณียกเว้นการเปลี่ยนแปลงทั้งเส้นทาง การทดสอบ ที่ไม่สามารถตรวจสอบได้

แนวทางการดำเนินงานต่อไปในอนาคตคือต้องการที่จะระบุสินทรัพย์ร่วมระหว่างผลิตภัณฑ์ตั้งแต่สองผลิตภัณฑ์ขึ้นไป รวมถึงการครอบคลุมสายผลิตภัณฑ์ซอฟต์แวร์

เอกสารอ้างอิง

- [1] P. Clements and L. Northrop, *Software product lines: practices and patterns*, Addison-Wesley Longman Publishing Co., Inc., 2001.
- [2] C.K. Roy and J.R. Cordy, "A Survey on Software Clone Detection Research," *SCHOOL OF COMPUTING TR 2007-541, QUEEN'S UNIVERSITY*, vol. 115, 2007.
- [3] R. Koschke, "Survey of Research on Software Clones", *Dagstuhl, Germany: Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI)*, Schloss Dagstuhl, Germany, 2007.
- [4] S. Bellon, R. Koschke, G. Antoniol, J. Krinke, and E. Merlo, "Comparison and Evaluation of Clone Detection Tools," *IEEE Transactions on Software Engineering*, vol. 33, 2007, pp. 591, 577.
- [5] Paul C.Jorgensen.2007. *–3rd ed Software testing:a craftsman approach*.Boca Raton New York. Auerbach Publications.
- [6] Tanit Reantragoon, "Identifying Common Asset Candidates in Software Product Line by Code Clone Detection" , *4th International Conference on Computer Research and Development*,2012.
- [7] Vijayanand Nagarajan, Rajiv Gupta, Matias Madou, Xiangyu Zhang, Bjorn De Sutter: "Matching Control Flow of Program Versions". *ICSM 2007*: 84-93
- [8] Miryung Kim, David Notkin. 2006. "Program Element Matching for Multi-Version Program Analyses". *MSR'06*.
- [9] A. Orso, N. Shi, and M. J. Harrold. "Scaling regression testing to large software systems". *SIGSOFT '04/FSE-12*, pages 241{251, 2004.
- [10] J. Laski and W. Szermer. "Identification of program modifications and its applications in software maintenance". *ICSM*, 1992.
- [11] L. A. Clarke, A. Podgurski, D. J. Richardson, and S. J. Zeil.. "A Formal Evaluation of Data Flow Path Selection Criteria". *IEEE Trans. Softw. Eng.* 15, 11 (November 1989), 1318-1332.
- [12] C. K. Roy and J. R. Cordy, "NICAD: Accurate detection of near-miss intentional clones using flexible pretty-printing and code normalization," in Program Comprehension, 2008. *ICPC 2008. The 16th IEEE International Conference on*, 2008, pp. 172–181.

ระบบสารสนเทศจัดการแปลงเพาะปลูกอ้อย Management Information System for Sugar Cane Garden

ณรรฐวรรณ พูลสน (Natthawan Phoonson)¹, ปวีรพรต คำคุณคำ (Pariwat Kumkunkum)² และ

อรรถพล หัตถะปณีตย์ (Autthapon Hutthapanit)³

สาขาวิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏอุดรธานี

398@udru.ac.th¹, porvp_@hotmail.co.th² and man_09_07_2534@hotmail.com³

บทคัดย่อ

ระบบสารสนเทศจัดการแปลงเพาะปลูกอ้อยเป็นการนำเทคโนโลยีด้านคอมพิวเตอร์พกพามาประยุกต์ใช้ในการเก็บข้อมูลแปลงเพาะปลูกอ้อยบนสมาร์ตโฟนที่ใช้ระบบปฏิบัติการแอนดรอยด์ เป็นโปรแกรมที่มีการพัฒนาสองสถาปัตยกรรมคือ ส่วนของโมบายแอปพลิเคชันและส่วนของเว็บแอปพลิเคชัน ส่วนโมบายแอปพลิเคชันเป็นระบบงานสำรวจและจัดเก็บพื้นที่เพาะปลูก มีเป้าหมายการทำงานเพื่อให้เจ้าหน้าที่ในภาคสนามเข้าถึงข้อมูลเชิงพื้นที่โดยการเรียกใช้บริการจาก Google Maps API และข้อมูลประกอบ มีการเชื่อมโยงกับฐานข้อมูลหลัก และอัปเดตข้อมูลที่สำรวจขึ้นเซิร์ฟเวอร์ได้ ส่วนที่สองระบบสารสนเทศสำหรับเจ้าหน้าที่ หัวหน้างาน และผู้บริหารเป็นบริการผ่านอินเทอร์เน็ต มีเป้าหมายเพื่อนำเสนอข้อมูลสารสนเทศเชิงพื้นที่ซึ่งได้จากดึงข้อมูลจากฐานข้อมูลหลักที่อยู่บนเซิร์ฟเวอร์ของโรงงานและสรุปรายงานเพื่อนำไปประเมินและวางแผน และส่งทีมงานไปแนะนำชาวไร่อ้อยในการเพาะปลูกต่อในขั้นต่อไป

คำสำคัญ: แปลงเพาะปลูกอ้อย โมบายแอปพลิเคชัน ระบบสารสนเทศภูมิศาสตร์, Google Maps API

Abstract

The aim of this study was to manage the sugar cane garden by using management information system. In this paper, the smart phone on android operating system was applied for data storage of sugar cane garden. The proposed method has developed two architectures which can use on the mobile application and web application.

The mobile application was used in survey and data storage of agricultural cultivation area for staff. The staff can manage the data in each area by using the Google Maps API which is linked with main database of system. The web application was used for chief and executive. The chief and executive can use management information system for evaluating and planning of agricultural cultivation in sugar cane garden efficiently. The results show that the proposed method performs very well which appropriate for evaluating and planning of agricultural cultivation in sugar cane garden.

Keywords: Sugar Cane Garden, Mobile Application, Geographic Information System, Google Maps API

1. บทนำ

ในอดีตการลงสำรวจแปลงเพาะปลูกอ้อยของทางโรงงานน้ำตาลจะส่งทีมสำรวจเข้าไปในพื้นที่แปลงเพาะปลูกอ้อยของเกษตรกรเพื่อจดบันทึกข้อมูลลงในกระดาษว่าอยู่ในช่วงไหนของการเพาะปลูก ขนาดของพื้นที่ในการเพาะปลูกมีกี่ไร่ มีลักษณะอย่างไร ติดกับสถานที่สำคัญอะไรบ้าง ทีมงานต้องอาศัยการวาดสเกลของพื้นที่ด้วยตนเอง ซึ่งไม่ค่อยแม่นยำ

เมื่อเทคโนโลยีเจริญขึ้นทุกครั้งที่พนักงานส่งเสริมชาวไร่ อ้อยลงพื้นที่เพื่อพบเกษตรกรในโครงการ พนักงานจะพกอุปกรณ์ GPS (Global Position System) เพื่อใช้กำหนดพิกัดของแปลงเพาะปลูกของเกษตรกรแต่ละราย และมีการจดบันทึกข้อมูลประกอบอื่นๆ ลงบนฟอร์มการจัดเก็บที่ทางบริษัทจัดทำขึ้น เมื่อกลับไปยังโรงงานน้ำตาลจะต้องใช้เวลาอีกหลายวันที่จะ

คือข้อมูลต่างๆ ที่ลงไปสำรวจ อาจมีข้อผิดพลาด ไม่ครบถ้วน จากการอ่านลายมือที่ไม่ชัดเจน ถ้าเรามีเครื่องมือที่เป็นอุปกรณ์พกพาที่สะดวกสามารถนำไปที่ไร้อ้อยแล้วสามารถแสดงพิกัดตำแหน่งที่อยู่ขณะนั้น ทำให้ได้ภาพถ่ายดาวเทียมที่ตรงกับสถานที่ก่อนนำภาพของไร้อ้อยในโครงการทั้งหมดมาต่อเรียงเข้าด้วยกันเป็นแผนภาพขนาดใหญ่ออกมา และทำการประเมินสภาพ วางแผนและส่งทีมไปแนะนำชาวไร่ในการเพาะปลูก

ในปัจจุบันระบบเทคโนโลยีสารสนเทศได้มีการประยุกต์ใช้แอปพลิเคชันดังเช่น ระบบสารสนเทศภูมิศาสตร์ด้านสิ่งแวดล้อม [2] การพัฒนาระบบแผนที่อาชีวกรรมกรณีสึกษา การโครงการมรดกในพื้นที่อำเภอเมือง จังหวัดนครปฐม [1], การประยุกต์ใช้ Google Map API ในการเฝ้าระวังปัญหาสุขภาพจิต [3], และ ระบบนำทางภายในมหาวิทยาลัยขอนแก่นบนระบบ Android [4] แต่อย่างไรก็ตามในการจัดการเรื่องอ้อยเพื่อส่งเสริมการเกษตรยังไม่มีความคืบหน้า ดังนั้นงานวิจัยนี้จึงได้ดำเนินการจัดทำระบบสารสนเทศจัดการแปลงเพาะปลูกอ้อยเพื่อส่งเสริม จัดการ และสร้างมูลค่า เพื่อพัฒนาระบบสารสนเทศทางการเกษตร

จากที่มาและปัญหาดังกล่าวข้างต้นทีมงานจึงคิดพัฒนาระบบสารสนเทศจัดการแปลงเพาะปลูกอ้อยบนอุปกรณ์ที่สะดวกในการพกพาไปสำรวจพื้นที่เช่น สมาร์ทโฟน ประกอบด้วย 2 ส่วน ส่วนที่หนึ่งเป็นระบบงานสำรวจและจัดเก็บพื้นที่เพาะปลูก มีเป้าหมายการทำงานเพื่อให้เจ้าหน้าที่ในภาคสนามเข้าถึงข้อมูลเชิงพื้นที่และข้อมูลประกอบ มีการเชื่อมโยงกับฐานข้อมูลหลัก และอัพโหลดข้อมูลขึ้นเซิร์ฟเวอร์ได้ ส่วนที่สองระบบที่เป็นบริการผ่านอินเทอร์เน็ต มีเป้าหมายเพื่อนำเสนอข้อมูลสารสนเทศเชิงพื้นที่ซึ่งได้จากดึงข้อมูลจากฐานข้อมูลหลักที่อยู่บนเซิร์ฟเวอร์ของโรงงานและสรุปรายงานแก่ผู้บริหารในรูปแบบของข้อความ กราฟ อื่นๆ เพื่อนำไปประเมินและวางแผนต่อไป

2. ระบบสารสนเทศภูมิศาสตร์

ระบบสารสนเทศภูมิศาสตร์ คือ กระบวนการทำงานเกี่ยวกับข้อมูลในเชิงพื้นที่ด้วยระบบคอมพิวเตอร์ ที่ใช้กำหนดข้อมูลและสารสนเทศ ที่มีความสัมพันธ์กับตำแหน่งในเชิงพื้นที่ องค์ประกอบหลักของระบบ GIS จัดแบ่งออกเป็น 5 ส่วนใหญ่

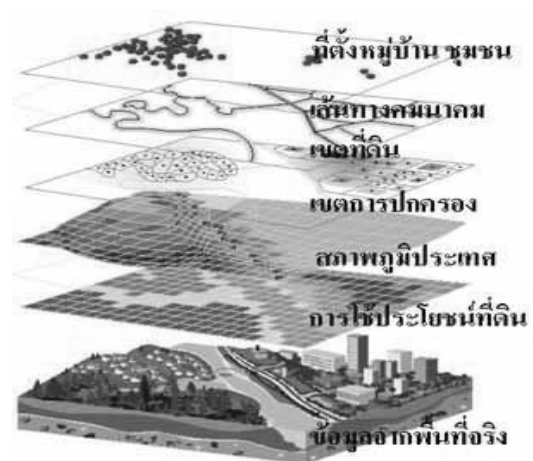
คือ อุปกรณ์คอมพิวเตอร์ (hardware) โปรแกรม (software) ขั้นตอนการทำงาน (methods) ข้อมูล (data) และบุคลากร (people) โดยมีรายละเอียดของแต่ละองค์ประกอบดังภาพที่ 1



ภาพที่ 1: องค์ประกอบของระบบสารสนเทศภูมิศาสตร์
(ที่มาจาก <http://yingpew103.wordpress.com>)

ข้อมูลเชิงพื้นที่ (Spatial Data) เป็นข้อมูลที่แสดงตำแหน่งที่ตั้งทางภูมิศาสตร์ (Geo-Reference Data) ของรูปลักษณะของพื้นที่ (Graphic Feature) ซึ่งมีอยู่ 2 แบบ คือ ข้อมูลที่แสดงทิศทาง (Vector Data) และข้อมูลที่แสดงเป็นตารางกริด (Raster Data) โดยข้อมูลที่มีทิศทาง ประกอบด้วยลักษณะ 3 อย่าง คือ

- ข้อมูลจุด (Point) เช่น ที่ตั้งหมู่บ้าน โรงเรียน หรือวัด
- ข้อมูลเส้น (Line) เช่น ถนน แม่น้ำ เป็นต้น
- ข้อมูลพื้นที่ หรือเส้นรอบรูป (Polygon) เช่น แหล่งน้ำผิวดิน แหล่งเพาะปลูกทางการเกษตร ดังภาพที่ 2



ภาพที่ 2: ลักษณะของข้อมูลทางภูมิศาสตร์
(ที่มาจาก <http://yingpew103.wordpress.com>)

3. Google Maps API

Google Maps API เป็นโปรแกรมที่ช่วยให้เราสามารถพัฒนาแอปพลิเคชันเพื่อแทรกตำแหน่งแผนที่เข้าไปในระบบ มีขีดความสามารถกว้างขวางเน้นในด้านการนำเสนอข้อมูลแผนที่ในลักษณะหมุดปัก (Push pin / Place marker) ซึ่งสามารถกำหนดให้แสดงข้อมูลประกอบแผนที่เมื่อผู้ใช้คลิกที่ตัว push pin /marker นั้นๆ หรือองค์แผนที่แบบเส้น (Polyline) พื้นที่ (Polygon) และภาพ (Ground overlay) ถ้าจะใช้ Google Maps API มาเพื่อที่จะพัฒนาแอปพลิเคชันจะต้องมี Account Email ของ Google และทำการสมัครเปิดใช้งาน API ก่อน บริการ API ของ Google ประกอบด้วย

-Google Maps JavaScript API v3 - เขียนโปรแกรมแผนที่ให้บริการผ่านเว็บไซต์

-Google Maps SDK for iOS - เขียนโปรแกรมแผนที่บนระบบปฏิบัติการ iOS Apple

-Google Maps Android API v2 - เขียนโปรแกรมแผนที่บนระบบปฏิบัติการ Android

-Google Maps Image APIs - เขียนโปรแกรมแผนที่ให้บริการผ่านเว็บไซต์แบบรูปภาพ

-Google Places API - เขียนโปรแกรมแผนที่ค้นหาสถานที่ผ่านสมาร์ตโฟน

4. โหมบายแอปพลิเคชัน

แนวโน้มการใช้งานสมาร์ตโฟนมีขอดีเพิ่มขึ้นอย่างก้าวกระโดดซึ่งเป็นผลจากการพัฒนาโปรแกรมประยุกต์บนอุปกรณ์มือถือและเทคโนโลยีของตัวเครื่อง โดยแบ่งเป็น 2 ประเภทดังนี้

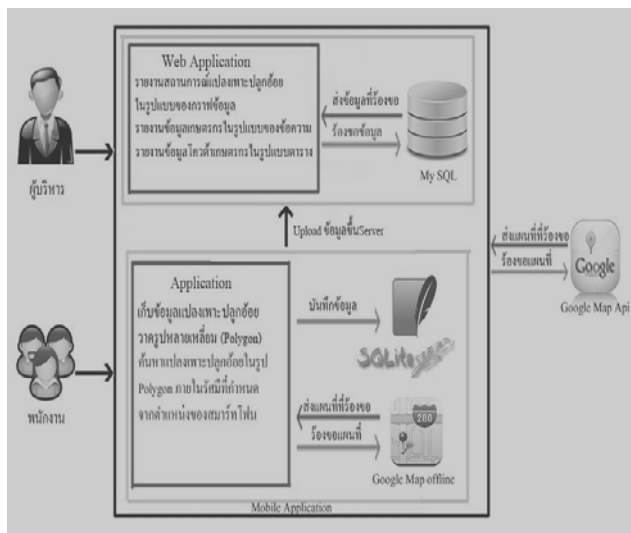
1) แอปพลิเคชันระบบ เป็นซอฟต์แวร์ระบบที่ใช้รองรับการใช้งานของแอปพลิเคชันต่าง ๆ ได้ ปัจจุบันนี้ระบบปฏิบัติการที่ได้รับความนิยมได้แก่ android, iOS, RIM และอื่นๆ ดังภาพที่ 3

ภาพที่ 4 กราฟแสดงการส่งอ้อยเข้าโรงงานเอราวัณ

การลงพื้นที่สำรวจของเจ้าหน้าที่พนักงานในการเก็บข้อมูลของแปลงเพาะปลูกอ้อยจะมีเครื่องมือในการเก็บข้อมูลคือเครื่อง GPS และกระดาษแบบฟอร์มการลงพื้นที่สำรวจ และจะเป็นแบบฟอร์มหลักที่ใช้ในการกรอกข้อมูลต่างๆของแปลงเพาะปลูกอ้อยเช่นหมายเลขเครื่อง GPS หมายเลขโคตดา ชื่อเกษตรกร ชื่อไร่ ที่ตั้งบ้านคำพิกัด X,Y พื้นที่แจ้งปลูก เอกสารสิทธิ์ สถานะครองที่ดิน การให้น้ำ สภาพพื้นที่ พันธุ์อ้อยที่ปลูก ประเภทอ้อย เป็นต้น และในแต่ละปีจะลงเก็บข้อมูลปีละ 3 ครั้ง

5.2 แนวคิดในการพัฒนาระบบงาน

ในการพัฒนาระบบสารสนเทศจัดการแปลงเพาะปลูกอ้อยจากข้อมูลที่ได้รวบรวมมานั้น มีการแบ่งการทำงานเป็น 2 โมดูลคือ โมดูลของการพัฒนาบนโมบายแอปพลิเคชันเพื่ออำนวยความสะดวกให้เจ้าหน้าที่ฝ่ายลงสำรวจข้อมูล และ โมดูลของการพัฒนาเว็บไซต์บนเครือข่ายอินเทอร์เน็ตเพื่อใช้จัดการข้อมูลพื้นฐานของฝ่ายดูแลและการสรุปรายงานเทสให้กับฝ่ายบริหารเพื่อดูภาพรวม สามารถสรุปองค์ประกอบที่เกี่ยวข้องและ โมดูลการทำงานหลักของระบบดังภาพที่ 5



ภาพที่ 5: ภาพรวมระบบ

ผู้ใช้งานระบบคือ ผู้บริหาร และ พนักงาน ผ่านทางเว็บเบราว์เซอร์และสมาร์ตโฟน ส่วนโมดูลการทำงานทั้งหมดจะประมวลผลทำงานบนเครื่องเซิร์ฟเวอร์ มีการร้องขอและ

ตอบสนองส่งผ่านข้อมูลระหว่างกันผ่านเครือข่ายอินเทอร์เน็ต และจัดเก็บข้อมูลเชิงพื้นที่ลงเก็บในฐานข้อมูลในเซิร์ฟเวอร์

5.3 ผลการดำเนินงาน

จากโมดูลหลักๆ (1) โมดูลของการพัฒนาบนโมบายแอปพลิเคชัน (2) โมดูลการพัฒนาในส่วนของเว็บแอปพลิเคชัน

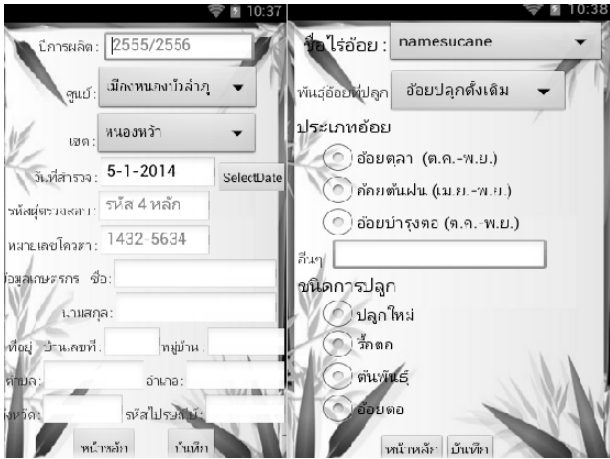
5.3.1 โมดูลของการพัฒนาบนโมบายแอปพลิเคชัน

ระบบงานได้พัฒนาการเก็บข้อมูลในรูปแบบออนไลน์และออฟไลน์ที่สามารถใช้งานบนระบบปฏิบัติการแอนดรอยด์เนื้อหาของ การจัดการสารสนเทศของพื้นที่แปลงเพาะปลูกอ้อย และการเก็บข้อมูลของระบบได้อย่างมีประสิทธิภาพ รวมทั้งการทำงานหน้าเว็บที่ใช้ในการเก็บข้อมูลของระบบให้ง่ายและสะดวกทั้งพนักงานและผู้บริหารของโรงงานเอราวัณ การรายงานผลการสำรวจของแต่ละศูนย์งานต่างๆได้โดยง่าย

รายการเมนูสำหรับเก็บข้อมูลแปลงเพาะปลูกอ้อยบน Mobile Application การเก็บข้อมูลสองรูปแบบในเครื่องเดียวคือ ฐานข้อมูล SQLite กรณีออฟไลน์ MySQL กรณีออนไลน์ การเก็บข้อมูลของเจ้าหน้าที่พนักงานซึ่งจะเก็บทั้งหมด 7 รายการ คือ ข้อมูลแปลงเพาะปลูกอ้อยที่สำรวจ, การทำสัญญาแปลงเพาะปลูกอ้อย, ข้อมูลสถานะแปลงเพาะปลูกอ้อย, ข้อมูลสถานะถือครองที่ดิน, ข้อมูลพันธุ์อ้อย, ข้อมูลสภาพพื้นที่, ข้อมูลแหล่งน้ำ ดังภาพที่ 6 และ 7

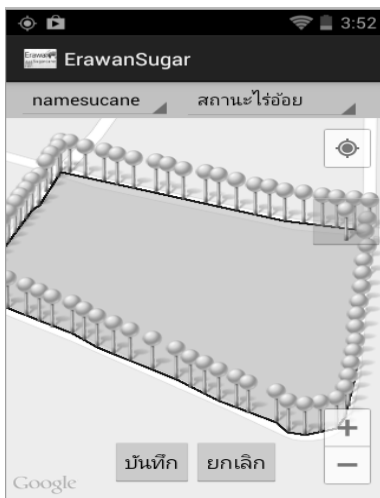


ภาพที่ 6: หน้าจอแรกของโปรแกรมบนมือถือ



ภาพที่ 7: หน้าจอรายการสำรวจข้อมูล

เก็บข้อมูลพื้นที่ของเกษตรกรไร่อ้อยโดยการปักหมุดในพื้นที่ โดยรอบของแปลงจะได้รูป Polygon ออกมาสามารถกำหนดสถานะไร่อ้อยได้โดยกำหนดสีสถานะ คือ สีเขียวหมายถึง ยังไม่ดำเนินการตัดอ้อย สีส้มหมายถึง กำลังดำเนินการตัด และสีแดงหมายถึงดำเนินการตัดเสร็จสิ้น

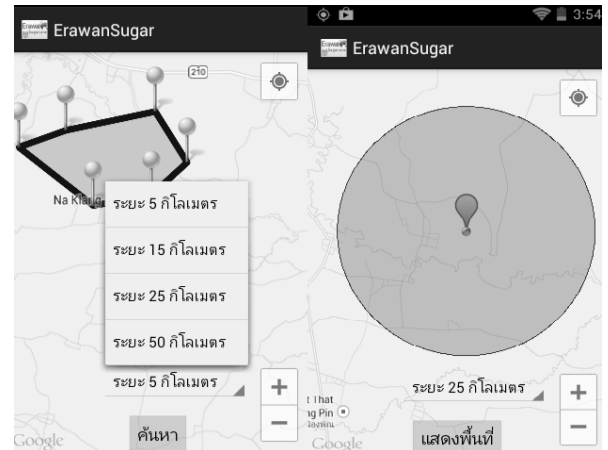


ภาพที่ 8: จอภาพแสดงการกำหนดจุดแปลงอ้อย

หน้าจอกำหนดรัศมีเป็นการกำหนดรัศมีจากตำแหน่งของ GPS ของเครื่องสมาร์ทโฟน กำหนดเพื่อให้แสดงพื้นที่ของแปลงเพาะปลูกของเกษตรกรแต่ละราย โดยมีรัศมี ดังนี้ ระยะ 5 กิโลเมตร, ระยะ 25 กิโลเมตร, ระยะ 50 กิโลเมตร

ในขณะที่เจ้าหน้าที่ออกไปเก็บข้อมูลยังพื้นที่แปลงปลูกอ้อยซึ่งหลายพื้นที่ที่ไม่สามารถเชื่อมต่อกับอินเทอร์เน็ตได้ ระบบจะมีโหมดในการจัดการให้ผู้ใช้งานสามารถใช้แผนที่ของ google

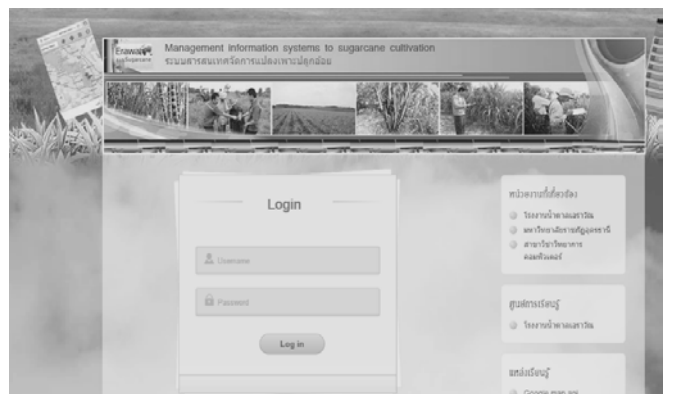
map ในตัวเครื่องสมาร์ทโฟนเก็บข้อมูลต่างๆ ไว้ในฐานข้อมูลที่มีการจัดเตรียมไว้ เมื่อเจ้าหน้าที่ทำงานเสร็จแล้วและกลับไปยังบริษัทหรือสถานที่ที่มีการเชื่อมต่อสัญญาณอินเทอร์เน็ต ระบบก็จะมีฟังก์ชันในการจัดการส่งข้อมูลในเครื่องสมาร์ทโฟนที่ไปเก็บมา จากแปลงเพาะปลูกอ้อยเข้าไปเก็บบนเซิร์ฟเวอร์ส่วนกลางต่อไป



ภาพที่ 9: หน้าจอการกำหนดรัศมีจากตำแหน่งปัจจุบัน

5.3.2 โมเดลของการพัฒนาบนเว็บแอปพลิเคชัน

ส่วนของเว็บแอปพลิเคชันจะแบ่งการทำงานออกเป็น 3 ส่วน คือ ผู้ดูแลระบบ, ผู้บริหาร, หัวหน้าศูนย์ โดยสิทธิ์การทำงานจะแตกต่างกัน ผู้ดูแลระบบจะทำหน้าที่ดูแลการทำงานข้อเว็บทั้งหมดรวมทั้งการจัดการข้อมูล ผู้บริหารจะทำหน้าที่ดูรายงานการทำงานของเจ้าหน้าที่พนักงานและข้อมูลในรูปแบบกราฟ และ ตารางข้อมูล หัวหน้าศูนย์ จะทำหน้าที่ในการจัดการกำลังคนในการลงพื้นที่สำรวจผ่านเว็บและทำหน้าที่ตรวจสอบความถูกต้องของข้อมูลการสำรวจ



ภาพที่ 10: จอภาพแรกของหน้าเว็บไซต์

ฟังก์ชันในการจัดการข้อมูลของหัวหน้าศูนย์เพื่อมอบหมายหน้าที่การทำงานให้พนักงานของแต่ละศูนย์ลงพื้นที่การสำรวจตามการมอบหมายงาน

รหัส	ชื่อเขต	พนักงาน
1236	กุฉีชัย	ผู้ดูแลสำรวจข้อมูล
1237	หนองบัวลำภู	ผู้ดูแลสำรวจข้อมูล
1242	กุมภวาปี	ผู้ดูแลสำรวจข้อมูล
1241	เขาสวนกวาง	ผู้ดูแลสำรวจข้อมูล

ภาพที่ 11: จอภาพการมอบหมายงานให้เจ้าหน้าที่ฝ่ายสำรวจ



ภาพที่ 12: ตัวอย่างการรายงานแปลงอ้อยที่ลงไปสำรวจ

6. บทสรุปและข้อเสนอแนะ

แอปพลิเคชันนี้พัฒนาขึ้นเพื่ออำนวยความสะดวกในการจัดเก็บข้อมูลเชิงบรรยายและข้อมูลเชิงพื้นที่ของแปลงเพาะปลูกของเกษตรกรไร่อ้อยและมีการสรุปผลให้ฝ่ายบริหารติดตามและประเมินผลการทำงานในภาพรวมได้ มีการพัฒนาโปรแกรมบนสองสถาปัตยกรรมคือ โมดูลการพัฒนาด้วยสถาปัตยกรรมโมบายแอปพลิเคชันเพื่ออำนวยความสะดวกให้เจ้าหน้าที่นำสมาร์ทโฟนไปจัดเก็บข้อมูลพื้นฐานของเกษตรกรและข้อมูลเชิงพื้นที่เพื่อกำหนดของเขตไร่อ้อยแต่ละแปลงของเกษตรกรที่ได้รับโควต้าไปในแต่ละปี เจ้าหน้าที่สามารถใช้งานระบบ ส่วนโมดูลที่สองพัฒนาบนสถาปัตยกรรมเว็บแอปพลิเคชันเพื่ออำนวยความสะดวกให้กับผู้ดูแลระบบ และหัวหน้างาน

แนวทางการพัฒนาต่อไปคือสามารถทำงานบนอุปกรณ์คอมพิวเตอร์พกพาได้ทั้งแท็บเล็ตคอมพิวเตอร์และสมาร์ทโฟน รวมถึงสามารถทำงานบนระบบปฏิบัติการไอโอเอส มีการใช้

อัลกอริทึมที่จะมาช่วยคำนวณจำนวนพื้นที่ของแปลงอ้อยที่ถูกต้องแม่นยำมากยิ่งขึ้น

7. กิตติกรรมประกาศ

ขอขอบคุณ สถาบันวิจัย มหาวิทยาลัยราชภัฏอุดรธานี ที่สนับสนุนทุนวิจัย และ ขอขอบคุณ โรงงานน้ำตาลเอราวัณที่เอื้อเฟื้อข้อมูลประกอบการศึกษาวิจัยครั้งนี้

เอกสารอ้างอิง

- [1] ชญา ณรงค์ฤทธิ์, “ระบบสารสนเทศภูมิศาสตร์ด้านสิ่งแวดล้อม,” *โครงการวิจัย การศึกษาพิเศษ มหาวิทยาลัยนเรศวร*, 2547.
- [2] นโกสินทร์ สุริยาจ, “การพัฒนาระบบแผนที่อาชญากรรมกรณีศึกษา การโจรกรรมรถในพื้นที่อำเภอเมือง จังหวัดนครปฐม,” *บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร*, 2553.
- [3] ปฐมพงษ์ ฉับพลัน และ จูนิภาพร เพชรแก้ว, “การประยุกต์ใช้ Google Map API ในการเฝ้าระวังปัญหาเสด็จ,” *สำนักวิชาสารสนเทศศาสตร์ มหาวิทยาลัยวลัยลักษณ์ นครศรีธรรมราช*, 2552.
- [4] ภาคภูมิ คล้ายบุญ และมาฆะพล หาปัญญา, “ระบบนำทางภายในมหาวิทยาลัยขอนแก่นบนระบบ Android,” *โครงการวิจัย การศึกษาพิเศษมหาวิทยาลัยขอนแก่น*, 2554.
- [5] สุระ พัฒนเกียรติ, “ระบบสารสนเทศภูมิศาสตร์เบื้องต้น,” [Online] (ม.ป.ป.). <http://www.sahavicha.com/UserFiles/File/GIS.doc> [อ้างเมื่อ 5 กรกฎาคม 2554].
- [6] อุไรวรรณ ศิริทอง, “ระบบระบุตำแหน่งบนพื้นโลก,” [Online] (ม.ป.ป.). <http://yinpew103.wordpress.com> [อ้างเมื่อ 5 มกราคม 2556].

เทคนิคการใช้วิดเจ็ตกับการควบคุมการจัดการการเปลี่ยนแปลงของซอฟต์แวร์ กรณีศึกษา:
ระบบเพื่อการจัดข้อมูลงานประกันคุณภาพการศึกษาของสถาบันพระบรมราชชนก PIEiS
**The technique using a widget to control the change management of software:
a case study of information systems for education quality assurance:
Praboromarajchanok Institute Executive Information System : PIEiS**

อริศา อ่อนเอื้อน (Athita Onuean)

สาขาวิชาวิศวกรรมซอฟต์แวร์ คณะวิทยาการสารสนเทศ มหาวิทยาลัยบูรพา

athitha@buu.ac.th

บทคัดย่อ

บทความนี้นำเสนอเทคนิคในการพัฒนาซอฟต์แวร์ที่มีการเปลี่ยนแปลงสูง โดยใช้วิดเจ็ต (Widget) เพื่อสามารถทำให้เกิดความยืดหยุ่นในการพัฒนาซอฟต์แวร์ ผู้ใช้งานสามารถปรับเปลี่ยนรูปแบบในการแสดงผลของข้อมูลได้ โดยได้ใช้ระบบเพื่อการจัดข้อมูลงานประกันคุณภาพการศึกษาของสถาบันพระบรมราชชนก หรือ Praboromarajchanok Institute Executive Information System : PIEiS เป็นกรณีศึกษาของซอฟต์แวร์ที่มีการเปลี่ยนแปลงสูง การวิจัยและพัฒนาซอฟต์แวร์นี้เพื่อเป็นการพัฒนาซอฟต์แวร์ต้นแบบ สำหรับเป็นข้อมูลสรุปของเกณฑ์ประกันคุณภาพการศึกษา สำหรับผู้บริหารเพื่อใช้ข้อมูลสารสนเทศในการตัดสินใจ

คำสำคัญ: วิดเจ็ต การพัฒนาระบบสารสนเทศ การควบคุมการเปลี่ยนแปลง ประกันคุณภาพการศึกษา

Abstract

This paper presents a technique to deal with software development with high rate of change. The technique involves using widget to enable flexibility in software development. With this approach, users can easily modify data visualization viewing on their own. We assure the concept by developing an information system prototype for education quality assurance: a case study of Praboromarajchanok Institute Executive Information System: PIEiS. The software provides the summary of the

education quality assurance criteria for executive's decision making.

Keyword: Widget, Software Development, Change Management, quality assurance of educational.

1. บทนำ

ในกระบวนการพัฒนาซอฟต์แวร์ตามหลักของวิศวกรรมซอฟต์แวร์ ผู้บริการโครงการจำเป็นต้องบริหารจัดการกับความเสี่ยงต่าง ๆ [1] เช่น ด้านงบประมาณ ด้านเวลา และด้านคน ดังนั้นสิ่งที่มีผลทำให้โครงการล้มเหลวหรือซอฟต์แวร์ทำไม่เสร็จตามกำหนดเวลาจึงเกิดได้จากหลายปัจจัย และหนึ่งในนั้นคือ ความเสี่ยงในเรื่องของความเปลี่ยนแปลงของซอฟต์แวร์ ผู้บริการโครงการซอฟต์แวร์จึงต้องรู้จักที่จะควบคุมจัดการการเปลี่ยนแปลงของซอฟต์แวร์ (Change Management) [2] ในบทความนี้ได้นำเสนอเทคนิคในการพัฒนาซอฟต์แวร์โดยประยุกต์ใช้แนวคิดวิดเจ็ต (Widget) ในการออกแบบและพัฒนาซอฟต์แวร์ระบบเพื่อการจัดข้อมูลงานประกันคุณภาพการศึกษาของสถาบันพระบรมราชชนก (PIEiS) ซึ่งเป็นซอฟต์แวร์ที่เป็นลักษณะการสรุปข้อมูลจากซอฟต์แวร์อื่น เพื่อใช้สำหรับเป็นข้อมูลรายงานตามเกณฑ์ตัวบ่งชี้ของงานประกันคุณภาพการศึกษา ซึ่งเกณฑ์ตัวบ่งชี้ดังกล่าวมีการเปลี่ยนแปลงเงื่อนไขต่าง ๆ บ่อยครั้ง จึงมีผลต่อการพัฒนาซอฟต์แวร์ โดยในบทความนี้จะประกอบไปด้วยส่วนแรกที่เป็นทฤษฎีและงานวิจัยที่เกี่ยวข้อง ส่วนที่สองจะเป็นการใช้วิดเจ็ตเพื่อควบคุมการจัดการการเปลี่ยนแปลง และส่วนที่สามจะเป็นการอธิบายการทำงานพื้นฐานของระบบ PIEiS

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 วิดเจ็ต (Widget)

ในปัจจุบันเว็บแอปพลิเคชันมีการพัฒนาเทคโนโลยีที่เป็นส่วนขยาย (Plugin Mechanisms) ซึ่งเป็นส่วนขยายของซอฟต์แวร์ โดยสามารถที่จะปรับเปลี่ยนข้อมูลภายในโดยผู้ใช้งาน ซึ่งเรียกว่า วิดเจ็ต (Widget) ตัวอย่างของซอฟต์แวร์โอเพนซอร์สที่พัฒนาในรูปแบบนี้ เช่น 1.) ซอฟต์แวร์ในกลุ่มการจัดการการเรียนรู้ (LMS) เช่น มูดเคิล (Moodle) แบล็กบอร์ด (Blackboard) 2.) ซอฟต์แวร์ในกลุ่มบล็อก เช่น เวิร์ดเพลส (Wordpress) 3.) ซอฟต์แวร์ในกลุ่มซอฟต์แวร์เพื่อสังคม เช่น เฟสบุ๊ก (Facebook) เอจ (Elgg) นิง (Ning) และ 4.) ซอฟต์แวร์ในกลุ่มที่เป็นระบบปฏิบัติการ เช่น แอปเปิ้ล เดสทอปบอร์ด (Apple Dashboard) วินโดวส์ไซด์บาร์ (Windows Sidebar) วิดเจ็ตของแอนดรอย (Android Widget) เป็นต้น ในแต่ละกลุ่มของซอฟต์แวร์ที่กล่าวมานั้น มี API ที่แตกต่างกัน ดังนั้นในการออกแบบและพัฒนาวิดเจ็ต ย่อมมีการออกแบบที่แตกต่างกันในแต่ละสภาพแวดล้อมของระบบอีกด้วย

ซึ่งจากเหตุผลข้างต้นจึงได้เกิดมาตรฐานเพื่อการสร้างวิดเจ็ตในแต่ละแพลตฟอร์มขึ้น ซึ่ง Google Gadgets and OpenSocial เป็นหนึ่งในกลุ่มที่จะพยายามสร้างเครื่องมือแบบวิดเจ็ตที่สามารถทำงานได้เป็นมาตรฐาน โดยใช้จาวาสคริปต์ APIs ควบคู่กับเว็บเซอร์วิสโดยใช้เทคนิคแบบ REST แต่การใช้ตามมาตรฐานของ Google Gadgets and OpenSocial มีความซับซ้อน และจำเป็นที่ผู้ดูแลเครื่องแม่ข่ายจะต้องปรับแต่งระบบปฏิบัติการให้รองรับกับสภาพแวดล้อมให้สามารถใช้เว็บเซอร์วิสได้อีกด้วย

นอกจากนี้ยังมีแนวคิดอื่น ๆ นั่นคือมาตรฐานของวิดเจ็ตของ W3C [3] และเป็นมาตรฐานเปิด (Open Standard) ที่มีการพัฒนาเครื่องมือที่จะประสานระหว่างแพลตฟอร์มของ แอปเปิ้ล ไมโครซอฟต์ ยาฮู โนเกียและโอเปรา โดยมีเป้าหมายในการมุ่งเน้นที่วิดเจ็ตที่พัฒนาบนเว็บ ที่ใช้ภาษาเอสทีเอ็มแอล จาวาสคริปต์ และซีเอสเอส เป็นต้น [4]

2.2 CodeIgniter : PHP Framework

CodeIgniter พัฒนาขึ้นโดย Rick Ellis เป็นเฟรมเวิร์คที่นำมาใช้ในการพัฒนาการเขียนเว็บแอปพลิเคชันด้วยภาษา PHP

อาศัยหลักการของ Model-View-Controller (MVC) ซึ่งสถาปัตยกรรมนี้แบ่งส่วนการพัฒนาออกเป็นสามส่วนอย่างชัดเจน คือ ส่วนการติดต่อผู้ใช้ (View) ส่วนการควบคุมการดำเนินงานไปของโปรแกรม (Controller) และส่วนการจัดการฐานข้อมูล (Model) ซึ่งการออกแบบสถาปัตยกรรมของระบบเช่นนี้ ทำให้ง่ายต่อการปรับปรุงแก้ไขหรือพัฒนาต่อยอด และมีความสอดคล้องกับการพัฒนาโปรแกรมเชิงวัตถุ (Object Oriented Programming) อย่างลงตัว โดยการนำเอากรอบงานของ CodeIgniter มาใช้เป็นแบบแผนในการพัฒนาระบบ นอกจากประโยชน์ในการนำมาใช้เป็นกรอบงานแล้ว CodeIgniter ยังเสริมความสามารถด้วยฟังก์ชันการทำงานพื้นฐานที่ผ่านการทดสอบจากกลุ่มผู้ใช้ทั่วโลก ทำให้ผู้พัฒนามีความมั่นใจในการนำไปใช้ระดับหนึ่ง

2.3 เว็บเซอร์วิส (Web Services)

เว็บเซอร์วิสคือซอฟต์แวร์ที่พัฒนาขึ้นเพื่อให้บริการแลกเปลี่ยนข้อมูลระหว่างเครื่องคอมพิวเตอร์ผ่านระบบเครือข่ายอินเทอร์เน็ตด้วยการร้องขอบริการเว็บเซอร์วิสตามสถาปัตยกรรม Representational State Transfer (REST) โดยใช้ภาษามาตรฐานที่เรียกว่า Extensible Markup Language (XML) ทำให้เครื่องคอมพิวเตอร์สามารถเรียกใช้งานเว็บเซอร์วิสและสื่อสารกันได้แม้ว่าจะมีแพลตฟอร์มหรือพัฒนามาจากภาษาที่แตกต่างกัน

1. SOAP (Simple Object Access Protocol)

เป็นโพรโทคอลที่มีโครงสร้างพื้นฐานจากภาษา XML ใช้เป็นสื่อกลางในการแลกเปลี่ยนข้อมูล UDDI (Universal Description, Discovery and Integration) เป็นมาตรฐานที่ระบุที่อยู่ของเว็บเซอร์วิสที่มีการลงทะเบียนไว้ เพื่อให้ผู้ใช้งานสามารถเข้ามค้นหาบริการต่าง ๆ ที่ได้ลงทะเบียนไว้

2. REST (Representational State Transfer)

เป็นสถาปัตยกรรมรูปแบบหนึ่งโดยอาศัยแนวคิดคือ สามารถทำงานบนโพรโทคอล HTTP และรับส่งข้อมูลในรูปแบบต่างๆ เช่น XML, XHTML เป็นต้น REST มีการแลกเปลี่ยนข้อมูลโดยใช้เมทอดเก็ท (Method GET) เพื่อดึงข้อมูล และ เมทอดพุท (Method PUT) เพื่อปรับปรุงข้อมูล

2.4 สถาปัตยกรรม REST

เว็บเซอร์วิสที่ใช้สถาปัตยกรรมแบบ REST (ในบางครั้ง

เรียกว่า RESTful) [5] เป็นรูปแบบของสถาปัตยกรรมทางซอฟต์แวร์สำหรับการนำเสนอข้อมูลต่าง ๆ โดยไม่ได้ถูกจำกัดให้ใช้งานโปรโตคอลใด ๆ แต่โดยทั่วไปมักจะใช้กับโปรโตคอล HTTP โดยรูปแบบการทำงานนั้นจะเป็นลักษณะการทำงานแบบร้องขอและแบบตอบกลับ (Request and Respond) โดยข้อมูลที่มีการรับส่งสามารถเป็นข้อมูลชนิดใดก็ได้ตามต้องการ เช่น XML, HTML, JSON เป็นต้น เว็บเซอร์วิสที่ให้บริการแบบ REST ซึ่งในบางครั้งเรียก RESTful จะเน้นที่ผลลัพธ์ของการให้บริการ เช่น cURL (ซึ่งได้กล่าวไว้ในส่วนถัดไป) ก็เป็นอีกรูปแบบหนึ่งที่ได้มีการใช้เทคนิคนี้ในการเขียนโปรแกรม โดยใช้วิธีการกำหนดที่อยู่ของทรัพยากรในระบบผ่านทาง URL ดังนั้นทุกครั้งที่มีการร้องขอบริการจะต้องมีการกำหนด URL เพื่อใช้ในการเรียกข้อมูลและทำงานร่วมกับ HTTP Method ต่างๆ

1. cURL คือ คำสั่งที่ใช้ในการรับส่งข้อมูล ซึ่งจะทำงานร่วมกับ URL โดยสนับสนุน FILE, FILES, FTP, FTPS, GOPHER, HTTP, HTTPS, IMAP, IMAPS, LDAP, LDAPS, POP3, POP3S, RTMP, RTSP, SCP, SFTP, SMTP, SMTPS, TELNET และ TFTP อีกทั้งยังสนับสนุน SSL, HTTP POST, HTTP PUT และเป็นฟังก์ชันที่ถูกใส่เข้ามาใน php ตั้งแต่ PHP 4.0.2. เป็นฟังก์ชันสำหรับใช้ในการติดต่อสื่อสารกับ server

2. การทำงานของ cURL จะใช้สถาปัตยกรรม REST ที่เน้นการใช้กระบวนการทำงานของ HTTP (POST, GET, PUT และ DELETE) ให้ถูกต้องและเหมาะสม การกำหนดรูปแบบของ URL ที่ใช้ในการรับส่งข้อมูล จะมีลักษณะคล้ายโครงสร้างของไคลเอนต์ โดยการใช้ URL เป็นตัวชี้ไปยังวัตถุ หรือที่อยู่ของบริการต่างๆ และสามารถกำหนดกระบวนการของ HTTP ไปพร้อมกับการร้องขอความต้องการ เพื่อให้แม่ข่ายเว็บเซอร์วิสดำเนินการส่งหรือรับทรัพยากรต่อไป

2.5 งานประกันคุณภาพการศึกษา

พระราชบัญญัติการศึกษาแห่งชาติ พ.ศ. 2542 แก้ไขเพิ่มเติม (ฉบับที่ 2) พ.ศ. 2545 ได้กำหนดจุดมุ่งหมายและหลักการของการจัดการศึกษาที่มุ่งเน้นคุณภาพและมาตรฐาน ซึ่งประกอบไปด้วย ระบบประกันคุณภาพภายในและระบบประกันคุณภาพภายนอก โดยระบบประกันคุณภาพภายในนั้นจะเป็นการสร้างระบบและกลไกในการตรวจสอบและประเมินการดำเนินงาน

ของสถานศึกษาให้เป็นไปตามนโยบายของสถานศึกษานั้น ๆ ซึ่งทางสำนักงานคณะกรรมการการอุดมศึกษา (สกอ.) ได้มีหน้าที่ในการกำหนดมาตรฐานที่เกี่ยวข้องกับการศึกษา [6] เพื่อให้สถานศึกษาได้ใช้มาตรฐานนี้เป็นตัวกำหนดระบบประกันคุณภาพการศึกษาภายในหน่วยงาน

ในส่วนของระบบประกันคุณภาพภายนอกนั้น เป็นการประเมินคุณภาพการจัดการศึกษาเพื่อให้มีการติดตามและตรวจสอบคุณภาพและมาตรฐานของสถานศึกษา ซึ่งประเมินโดย “สำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์การมหาชน)” หรือเรียกชื่อย่อว่า สมศ. [7] และนอกเหนือจากมาตรฐานทั้งสองมาตรฐานดังกล่าวข้างต้น ยังมีมาตรฐานของสภากาชาด ซึ่งเป็นเกณฑ์มาตรฐานเพื่อให้การรับรองสถาบันการศึกษาให้เป็นไปอย่างถูกต้องตามข้อบังคับของสภากาชาด ซึ่งเป็นกระบวนการในการพิจารณาการจัดการศึกษาที่สภากาชาดให้ความเห็นชอบให้จัดการศึกษา วิชาชีวะการพยาบาลและการผดุงครรภ์ [8] โดยวิธีการตรวจสอบจากรายงานการประเมินตนเอง การตรวจเยี่ยม และการประเมินคุณภาพสถานศึกษา

2.6 ซอฟต์แวร์ระบบสารสนเทศเพื่อการจัดข้อมูลงานประกันคุณภาพการศึกษาของสถาบันพระบรมราชชนก: Praboromarajchanok Institute Executive Information System (PIEiS)

ซอฟต์แวร์ระบบสารสนเทศเพื่อการจัดข้อมูลงานประกันคุณภาพการศึกษาของสถาบันพระบรมราชชนก หรือมีชื่อย่อว่า PIEiS เป็นระบบที่พัฒนาขึ้นมาเพื่อใช้เป็นการนำเสนอข้อมูลที่เกี่ยวข้องกับตัวบ่งชี้ของข้อมูลงานประกันคุณภาพการศึกษาใน 3 เกณฑ์ตัวบ่งชี้ ได้แก่ 1.) เกณฑ์ของสภากาชาด 2.) เกณฑ์ของ สกอ. 3.) เกณฑ์ของ สมศ. และ ตัวบ่งชี้เฉพาะของสถาบันพระบรมราชชนก ซอฟต์แวร์นี้เป็นระบบ Output System กล่าวคือเป็นระบบที่มีการสกัดข้อมูลจากระบบอื่น ๆ ของสถาบันพระบรมราชชนก มีความทำงานที่ยืดหยุ่น สามารถปรับแต่งค่าได้ (ในระดับหนึ่ง) และสามารถปรับเปลี่ยนการแสดงผลให้สอดคล้องของแต่ละตัวบ่งชี้ (Widgets Customization) ข้อมูลที่สกัดมาจาก PIEiS จะเกิดจากข้อมูลดิบของระบบสารสนเทศพื้นฐาน ภายใต้เงื่อนไขที่ผู้ใช้เป็นผู้กำหนด ข้อมูลที่ถูกนำเสนอใน PIEiS จะมีรูปแบบที่เข้าใจง่าย

โดยในภาพที่ 1 แสดงแผนผังระบบสารสนเทศที่เป็นรากฐานของระบบ PIEiS ได้แก่ ระบบบุคลากร ระบบสารบรรณ ระบบแผน ระบบ E-meeting ระบบกำกับงบประมาณ ระบบทะเบียนนักศึกษา ระบบ TQF ระบบภาระงานสอน และระบบอื่นๆ ซึ่งข้อมูลเหล่านี้จะถูกสกัดไปเป็นข้อมูลเชิงสรุปที่ ระบบ PIEiS



ภาพที่ 1: แผนผังระบบสารสนเทศที่เป็นรากฐานของระบบ

3. การใช้วิเจ็ดเพื่อควบคุมการจัดการการเปลี่ยนแปลง

แนวคิดการใช้วิเจ็ดเพื่อจัดการการเปลี่ยนแปลง เกิดจากการรวบรวมข้อมูลการพัฒนาซอฟต์แวร์ให้กับหน่วยงานสถาบันพระบรมราชชนกและวิทยาลัยในสังกัด ซึ่งห้องปฏิบัติการวิจัยวิศวกรรมระบบสารสนเทศ คณะวิทยาการสารสนเทศ มหาวิทยาลัยบูรพาได้ วิจัย ศึกษาและพัฒนาซอฟต์แวร์ต่างๆ มาอย่างต่อเนื่องเป็นเวลา 8 ปี พบว่าผู้ใช้มีความต้องการในการปรับเปลี่ยนซอฟต์แวร์ที่สูงมาก เนื่องมาจากบริบทในการทำงานของวิทยาลัยเหล่านั้นมีความแตกต่างกัน แต่ทางห้องปฏิบัติการวิจัยฯ จะต้องพัฒนาซอฟต์แวร์เพียง หนึ่งรูปแบบ เพื่อให้ใช้งานได้ทั้งในส่วนกลางและวิทยาลัย จึงต้องออกแบบซอฟต์แวร์ให้สามารถทำการปรับแต่ง (Configuration) ได้ทุกระดับของการทำงานในทุกซอฟต์แวร์ที่พัฒนาอยู่

ในส่วนของระบบ PIEiS ซึ่งเป็นระบบขาออก ซึ่งทำหน้าที่สกัดข้อมูลอื่นๆ จากระบบพื้นฐานด้านล่าง ดังนั้นในแต่ละหน้าจอของซอฟต์แวร์จึงได้ออกแบบเป็นเทมเพลต ซึ่งได้แนวคิดจากหน้าจอของโทรศัพท์มือถือสมาร์ตโฟน ที่ผู้ใช้สามารถเลือกเพิ่ม-ลดหน้าจอได้ และในแต่ละหน้าจอผู้ใช้สามารถเลือกใส่ไอคอน หรือวิเจ็ดที่ต้องการได้ จากแนวคิดดังกล่าวผู้วิจัยจึงได้วิเคราะห์ข้อมูลตัวบ่งชี้และทุกเกณฑ์ แล้วมาออกแบบให้อยู่ในรูปแบบของวิเจ็ด ข้อมูลที่ใช้งานบ่อยจะ

ออกแบบเป็นวิเจ็ดมาตรฐาน สามารถเลือกใช้ได้ในทุกเทมเพลต และวิเจ็ดเฉพาะจะเป็นตัวบ่งชี้ที่ต้องทำการสกัดข้อมูลจากระบบสารสนเทศต่าง ๆ ที่พัฒนาโดยห้องปฏิบัติการวิจัยฯ โดยจะแสดงในรูปแบบของการเลือกข้อมูล ทั้งแบบอัตโนมัติ หรือผู้ใช้เลือกข้อมูลด้วยตนเอง สามารถเลือกใช้วิเจ็ดเฉพาะในบางเทมเพลตที่เกี่ยวข้อง

ดังนั้นในการพัฒนาซอฟต์แวร์ต่อไปในอนาคตในกรณีที่ตัวบ่งชี้ในแต่ละเกณฑ์มีการเปลี่ยนแปลงเงื่อนไขของเกณฑ์ ผู้พัฒนาจะไม่จำเป็นต้องรี้อโครงสร้างของซอฟต์แวร์ แต่จะสามารถพัฒนาวิเจ็ดใหม่ ๆ เพิ่มเติมขึ้นมา เพื่อให้ผู้ใช้สามารถเลือกวิเจ็ดที่เป็นเกณฑ์ใหม่เหล่านั้น ไปใส่ในเทมเพลตได้ อีกทั้งยังสามารถที่จะเลือกทั้งเกณฑ์เก่าและเกณฑ์ใหม่การเปรียบเทียบผลคะแนนต่าง ๆ ได้อีกด้วย

4. การทำงานพื้นฐานของระบบ PIEiS

ในส่วนนี้จะกล่าวถึงแนวคิดของการพัฒนาระบบ PIEiS ฟังก์ชันการทำงานหลักของระบบ PIEiS ในส่วนแรกคือส่วนวิเจ็ดมาตรฐาน ส่วนที่สองคือส่วนของวิเจ็ดเฉพาะ และส่วนที่สามเป็นการนำเสนอรูปแบบของหน้าจอตัวบ่งชี้

4.1 แนวคิดในการพัฒนาระบบ PIEiS

1. เกิดจากความต้องการในการพัฒนาระบบสารสนเทศที่สามารถปรับแก้ไขข้อมูลได้ง่าย กล่าวคือ เมื่อมีการเปรียบเทียบข้อมูลตัวบ่งชี้ของการประกันคุณภาพการศึกษาจาก 3 ตัวบ่งชี้ พบว่ามีเกณฑ์ต่างๆ ที่ต้องการข้อมูลที่เหมือนกัน หรือใกล้เคียงกัน และเกณฑ์เหล่านี้มีการเปลี่ยนแปลงบ่อย แต่มีการเปลี่ยนแปลงที่ไม่มากนัก ดังนั้นหากพัฒนาซอฟต์แวร์แบบคอนกรีต นักพัฒนาจะต้องเข้าไปปรับแก้ซอร์สโค้ดในในส่วนของสูตร ผู้ใช้งานไม่สามารถปรับเองได้

2. รูปแบบการนำเสนอ แม้ว่าระบบ PIEiS จะสกัดข้อมูลมาจากระบบสารสนเทศต่างๆ แต่วิทยาลัยในสังกัดสถาบันพระบรมราชชนกจำนวน 38 แห่ง มีความต้องการในการนำเสนอข้อมูลที่แตกต่างกัน ยกตัวอย่างเช่น ตัวบ่งชี้ที่ 5 คุณสมบัติอาจารย์พยาบาลประจำ ของเกณฑ์สภาพพยาบาล บางวิทยาลัยต้องการนำเสนอข้อมูลกราฟก่อน บางวิทยาลัยต้องการแสดงข้อมูลตารางก่อน บางวิทยาลัยของการนำเสนอข้อมูลใน

รูปแบบการบรรยายเชิงพรรณาก่อน ดังนั้นผู้พัฒนาจึงต้องออกแบบระบบสารสนเทศ เพื่อให้ผู้ใช้ปรับแต่งเองได้

3. ระบบ PIEiS จะสามารถแสดงรายงานที่สามารถทำให้ผู้บริการกำกับ ติดตาม ความก้าวหน้าในการทำงานผ่านระบบสารสนเทศต่าง ๆ ของสถาบันได้อย่างสม่ำเสมอ

4.2 ฟังก์ชันการทำงานหลักของระบบ PIEiS

ระบบ PIEiS จะประกอบด้วยผู้ใช้ 5 ส่วนคือ 1.) เจ้าหน้าที่ที่ดูแลประกันคุณภาพการศึกษา 2.) รองผู้อำนวยการวิทยาลัย ที่ดูแลข้อมูลในแต่ละตัวบ่งชี้ 3.) ผู้อำนวยการวิทยาลัย 4.) ผู้ดูแลระบบระดับวิทยาลัย 5.) ผู้ดูแลระบบระดับส่วนกลาง โดยในภาพที่ 2 แสดงหน้าจอเครื่องมือของผู้ดูแลระบบในส่วนของวิทยาลัย ประกอบไปด้วย

1. การแสดงผล ซึ่งในส่วนของฟังก์ชันนี้เป็นการกำหนดรูปแบบเทมเพลตของแต่ละเกณฑ์ แยกเป็นตัวบ่งชี้ ซึ่งมีเทมเพลตให้เลือกจำนวน 4 เทมเพลต ดังแสดงในภาพที่ 3
2. การเข้าถึงตัวบ่งชี้ เป็นส่วนของการกำหนดการเข้าถึงข้อมูลของแต่ละตัวบ่งชี้ โดยจะต้องระบุชื่อของบุคลากรในการเข้าไปจัดการข้อมูลในแต่ละเกณฑ์ แต่ละตัวบ่งชี้
3. เป็นส่วนปรับแต่งการเชื่อมประสานข้อมูลระหว่างระบบ PIEiS กับระบบสารสนเทศอื่น ๆ เช่น ระบบบุคลากร ระบบภาระงานสอน ระบบ TQF ระบบทะเบียนนักศึกษา ผลการประเมินอาจารย์ ระบบวิจัย เป็นต้น

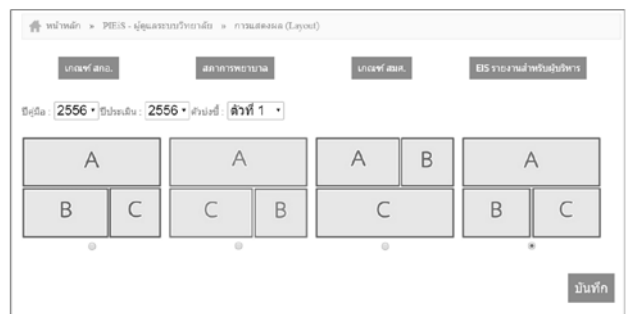
4.3 ปรับเปลี่ยนการแสดงผลให้สอดคล้องของแต่ละตัวบ่งชี้ (Widgets Customization)

ซอฟต์แวร์ระบบ PIEiS ประกอบไปด้วยวิดเจ็ตมาตรฐาน และวิดเจ็ตเฉพาะ โดยวิดเจ็ตมาตรฐานประกอบไปด้วย แสดงภาพกิจกรรม แสดงคำสำคัญ ประเมินตนเอง นำเข้าไฟล์ เอกสารแนบ ข้อความหลายบรรทัด ข้อความแบบแท็บ ข้อมูลลิงค์ไปยังเว็บไซต์ กราฟแบบวงกลม และกราฟแบบแท่ง โดยผู้ใช้สามารถที่จะเลือกวิดเจ็ตมาตรฐาน และกำหนดลงไป ในเทมเพลตตัวบ่งชี้ในแต่ละเกณฑ์ ในภาพที่ 4 แสดงหน้าต่างวิดเจ็ตทางด้านขวาของรูป ซึ่งจะมียึดมาตรฐาน โดยผู้ใช้สามารถเลือกวิดเจ็ตเหล่านั้น โดยทำการลากมาวางที่ ตัวบ่งชี้ทางด้านซ้าย ซึ่งเป็นกรอบเส้นประ ในมาตรฐานของตัวชี้วัดที่ 5 ใช้เทมเพลตแบบ C มีวิดเจ็ตแสดงตารางจำนวนอาจารย์พยาบาลที่ผ่านตามเกณฑ์ได้ มีวิดเจ็ตข้อความแบบหลายบรรทัด และวิด

เจ็ตเกณฑ์มาตรฐานของตัวบ่งชี้ โดยในการปรับเปลี่ยนตำแหน่งของวิดเจ็ต จะใช้เทคนิค Drag และ Drop หรือการลากวาง โดยผู้ใช้ในแต่ละวิทยาลัยสามารถที่จะปรับตำแหน่งของวิดเจ็ตตามที่ต้องการได้



ภาพที่ 2: เครื่องมือของผู้ดูแลระบบในส่วนของวิทยาลัย



ภาพที่ 3: เทมเพลตตัวบ่งชี้ในแต่ละเกณฑ์



ภาพที่ 4 : การแทรกวิดเจ็ตในเทมเพลต

4.4 ปรับเปลี่ยนข้อมูลให้สอดคล้องของแต่ละตัวบ่งชี้

นอกจากวิดเจ็ตมาตรฐานระบบ PIEiS ยังประกอบไปด้วยวิดเจ็ตเฉพาะ สำหรับในบางตัวบ่งชี้ โดยระบบ PIEiS จะไปเลือกข้อมูลโดยการเชื่อมประสานกับระบบอื่น ๆ จากนั้นผู้ใช้จะเลือกข้อมูลเพื่อใช้ในการตัดสินใจ ตัวอย่างของวิดเจ็ตมาตรฐานของตัวบ่งชี้ที่ 5 คุณสมบัตินักพยาบาลประจำ ซึ่งการที่ได้ออกแบบซอฟต์แวร์ ให้ผู้ใช้ตัดสินใจเลือกเอง เพราะเงื่อนไขของเกณฑ์ในการนับมีการเปลี่ยนแปลงบ่อยครั้งในทุกปี ดังนั้น

หากทำซอฟต์แวร์ให้ทำแบบอัตโนมัติ ถ้าเกณฑ์มีการเปลี่ยนแปลง ผู้พัฒนาจะต้องมาแก้ไขซอฟต์แวร์ในทุกครั้ง แต่ถ้าออกแบบวิเคาต์เฉพาะนี้ให้เป็ฯข้อมูลเพื่อให้ผู้ใช้เลือก และตัดสินใจเอง ก็จะสามารไปประยุกต์ใช้กับเกณฑ์ข้ออื่น ๆ ได้ อีก เช่น เงื่อนไขของการนับอาจารย์พยาบาลประจำ ได้แก่ ระยะเวลาที่ทำงาน วุฒิการศึกษาระดับปริญญาโทที่ต้องจบทาง สาขาพยาบาลศาสตร์ บัตรประกอบวิชาชีพ บัตรสภาพยาบาลที่ ต้องไม่หมดอายุ กรณีที่มีการโอนย้ายมาจากวิทยาลัยอื่น ต้องจัด กลุ่มระดับคะแนน เป็นต้น ดังแสดงในภาพที่ 5 คือตัวอย่างของ วิเคาต์เฉพาะที่เป็นการคำนวณร้อยละของจำนวนอาจารย์ พยาบาลประจำที่ได้ตามเกณฑ์



ภาพที่ 5: วิเคาต์เฉพาะที่เกี่ยวข้องกับข้อมูลบุคลากร

<input type="checkbox"/> ข้อมูลวุฒิการศึกษาสูงสุด ทั้งหมด ไม่เลือกทั้งหมด	<input type="checkbox"/> ข้อมูลตำแหน่งบริหาร ทั้งหมด ไม่เลือกทั้งหมด	<input type="checkbox"/> ข้อมูลการไปอบรม/พัฒนาตนเอง ทั้งหมด ไม่เลือกทั้งหมด
<input type="checkbox"/> ข้อมูลอายุ	<input type="checkbox"/> ข้อมูลประเภทสายงาน ทั้งหมด ไม่เลือกทั้งหมด	<input type="checkbox"/> วันที่บรรจุเข้ารับราชการครั้งแรก
<input type="checkbox"/> ข้อมูลระดับตำแหน่ง ทั้งหมด ไม่เลือกทั้งหมด	<input type="checkbox"/> ข้อมูลการศึกษา ทั้งหมด ไม่เลือกทั้งหมด	<input type="checkbox"/> วันที่เริ่มปฏิบัติงานที่วิทยาลัย
		<input type="checkbox"/> การเป็นอาจารย์พยาบาลประจำ

ภาพที่ 6: การเลือกค้นหาข้อมูลบุคลากร

5. สรุป

บทความนี้เป็นกรนำเสนอเทคนิคการใช้วิเคาต์กับการควบคุมการจัดการการเปลี่ยนแปลงของซอฟต์แวร์ ที่ได้ ออกแบบมาเพื่อช่วยในการจัดการความเปลี่ยนแปลงที่อาจจะเกิดขึ้นในกระบวนการพัฒนาซอฟต์แวร์ โดยเฉพาะอย่างยิ่งในกรณีศึกษาซอฟต์แวร์ที่ระบบ PIEiS ที่เป็นซอฟต์แวร์เพื่อการจัด ข้อมูลงานประกันคุณภาพการศึกษาของสถาบันพระบรมราชชนก

ที่มีการเปลี่ยนแปลงซอฟต์แวร์ทั้งในส่วนที่เป็นเกณฑ์ตัวบ่งชี้ และอินเทอร์เฟซของซอฟต์แวร์สูง ปัจจุบันมี 7 วิทยาลัยที่เป็นหน่วยงานนำร่องในการในการใช้ระบบ PIEiS ในการตรวจรับ การประเมินประกันคุณภาพการศึกษาเกณฑ์ของสภาพยาบาล ประจำปีการศึกษา 2556 และในปี 2557 จะเริ่มใช้งานกับ วิทยาลัยอื่นๆ ที่มีความพร้อมต่อไป และคณะผู้วิจัยได้วางแผน ที่จะพัฒนาตัวบ่งชี้ในส่วนของ เกณฑ์ สกอ. และ สมศ. และคาด ว่าทางสถาบันพระบรมราชชนกจะสามารถใช้ในการตรวจรับ การประเมินในปีการศึกษา 2557 ต่อไป

เอกสารอ้างอิง

- [1] Adel M. Aladwani, "Change management strategies for successful ERP implementation", *Business Process Management Journal*, Vol. 7 Iss: 3, pp.266 – 275, 2001.
- [2] Kramer J. Magee J, " The evolving philosophers problem: dynamic change management ", *Software Engineering, IEEE Transactions on* Volume:16 , Issue: 11, pp.1293 – 1306, 1990.
- [3] Attwell, G., The personal learning environments – The future of eLearning? *eLearning Papers*, 2(1). [cited 2014 Feb 20]. Available from: <http://www.elearningeuropa.info/files/media/media11561.pdf>.
- [4] Packaged Web Apps (Widgets) - Packaging and XML Configuration (Second Edition). W3C Recommendation 27 November 2012. [cited 2014 Feb 26]. Available from: <http://www.w3.org/TR/widgets/>
- [5] พชรวร นุญชู, มยุรี เลิศเวชกุล. “การพิจารณาประสิทธิภาพการทำงานของเว็บเซอร์วิสที่ใช้สถาปัตยกรรม REST” *Proceedings of National Conference on Information Technology: NCIT2010*. pp 207-212, October 2010.
- [6] สำนักงานคณะกรรมการการอุดมศึกษา, “คู่มือการประกันคุณภาพการศึกษา”, สวัสดิการสำนักงานคณะกรรมการการอุดมศึกษา. กรุงเทพฯ 2554.
- [7] คู่มือการประเมินคุณภาพนอกรอบสาม (พ.ศ. 2554 - 2558) ระดับอุดมศึกษา ฉบับสถานศึกษา (แก้ไขเพิ่มเติม พฤศจิกายน พ.ศ. 2557) กรุงเทพฯ 2554.
- [8] คู่มือการรับรองสถาบันการศึกษาวิชาการพยาบาลและการผุ้จรรรภ. สภาการพยาบาล กรุงเทพฯ 2556.

ต้นแบบระบบสนับสนุนการตัดสินใจเกณฑ์การคัดแยก และจัดลำดับการให้บริการ
ผู้ป่วยฉุกเฉินตามเกณฑ์ กรณีศึกษาของสถาบันการแพทย์ฉุกเฉินแห่งชาติ
**Prototype of Decision Support System for Emergency Medical Triage Protocol
and Criteria Based Dispatch**
A Case Study of National Institute for Emergency Medicine

พิรดี เถลิรัตน์พงศ์ (Piradee Chalomprattanapong)¹ และ พงษ์พิสิฐ วุฒิชัยชูโชค (Pongpisit Wuttidittachotti)²

ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹ piradee.ch@rd.go.th, ² pongpisitw@kmutnb.ac.th

บทคัดย่อ

การรักษาให้ทันเวลาจำเป็นต่อชีวิตผู้ป่วย เมื่อการเจ็บป่วยฉุกเฉินเกิดขึ้น ผู้ป่วยฉุกเฉินหรือผู้แจ้งเหตุจะติดต่อมายังศูนย์รับแจ้งเหตุและสั่งการ จากนั้นศูนย์ฯ จะทำการวิเคราะห์และจัดส่งชุดปฏิบัติการไปยังจุดเกิดเหตุ โดยศูนย์ฯ ได้มีการจัดทำแอปพลิเคชันซึ่งเป็นแบบฟอร์มของเกณฑ์การคัดแยกและจัดลำดับการให้บริการผู้ป่วยฉุกเฉินตามเกณฑ์เท่านั้น โดยไม่มีการเก็บข้อมูลที่เกี่ยวข้องอื่น ๆ แอปพลิเคชันดังกล่าวมีความซับซ้อน หลายขั้นตอน ทำให้เกิดความยุ่งยากในการปฏิบัติงานจริง ไม่ทันต่อเหตุการณ์ จึงทำให้ผู้ปฏิบัติงานใช้ประสบการณ์ของตนเองในการวิเคราะห์ ซึ่งอาจทำให้ผลการวิเคราะห์ของผู้ปฏิบัติงานแต่ละคนได้ผลลัพธ์ที่คลาดเคลื่อนไปจากความเป็นจริง ดังนั้นผู้วิจัยจึงมีแนวคิดในการพัฒนาระบบโดยใช้ออนโทโลยีในการฐานความรู้ และ ใช้การแสดงผลแบบทรีแมพ (Tree map) เพื่อให้เกณฑ์ต่าง ๆ แสดงผลลัพธ์อยู่ภายในหน้าจอเดียวกัน ซึ่งจะช่วยให้ง่ายต่อการใช้งานและการตัดสินใจมีความรวดเร็วขึ้น

คำสำคัญ: ระบบสนับสนุนการตัดสินใจ ออนโทโลยี ทรีแมพ

Abstract

Treatment time is critical to the life of patients. When an emergency occurs, emergency patients or others related to patients will call to a dispatch center. Next, a dispatch center officer will analyze the case and deliver a suitable rescue team to the target. The dispatch center provided a windows application which is the basis

for Emergency Medical Triage Protocol and Criteria Based Dispatch only, no other associated information are kept. The application has complicated processes and multi-steps which cause the practical difficulties and do not respond in time. Thus, the officer uses their own experiences in the analysis process which leads to improper results deviated from reality. Therefore, we propose a prototype consisting of 2 techniques: 1) Ontology technology is utilized to create the knowledge base 2) Treemap technology is used to illustrate the various criteria results in a single screen. The prototype will allow an easier use and a quicker decision.

Keywords: Decision Support System, Ontology, TreeMap

1. บทนำ

การเจ็บป่วยฉุกเฉิน โดยไม่ได้รับการรักษาให้ทันเวลาถือเป็นความเสี่ยงต่อชีวิตผู้ป่วย ผู้แจ้งเหตุจะประสานมายังศูนย์รับแจ้งเหตุและสั่งการ [1] จากนั้นจึงจัดส่งชุดปฏิบัติการมารับผู้ป่วยฉุกเฉินไปรักษาที่สถานพยาบาลต่อไป

สถาบันการแพทย์ฉุกเฉินแห่งชาติได้มีการจัดตั้ง ศูนย์รับแจ้งเหตุและสั่งการ [1] ให้บริการตลอด 24 ชั่วโมง สายด่วน 1669 และจัดทำเกณฑ์การคัดแยกและจัดลำดับการให้บริการผู้ป่วยฉุกเฉินตามเกณฑ์ในรูปแบบของแอปพลิเคชันติดตั้งไว้ที่เครื่อง Client ในทุกศูนย์ฯ ซึ่งปัญหาที่เกิดขึ้น คือ แอปพลิเคชันที่พัฒนานั้นเป็นแค่แบบฟอร์มคำถามของเกณฑ์

การคัดแยกและจัดลำดับการให้บริการผู้ป่วยฉุกเฉินตามเกณฑ์เท่านั้น ไม่มีการเก็บข้อมูลที่เกี่ยวข้องอื่น ๆ แต่อย่างใด เช่น ข้อมูลสถานที่เกิดเหตุ ข้อมูลผู้แจ้งเหตุ เบอร์โทรศัพท์ที่สามารถติดต่อได้ เป็นต้น ทำให้ไม่สามารถนำข้อมูลไปวิเคราะห์ต่อได้ แอปพลิเคชันดังกล่าวมีความซับซ้อน มีหลายขั้นตอน มีลักษณะเป็นต้นไม้ตัดสินใจ จะไม่สามารถข้ามขั้นตอนได้ ทำให้เกิดความยุ่งยากในการปฏิบัติงานจริง ไม่ทันต่อเหตุการณ์ จึงทำให้ผู้ปฏิบัติงานใช้ประสบการณ์ของตนเองในการวิเคราะห์ ซึ่งอาจทำให้ผลการวิเคราะห์ของผู้ปฏิบัติงานแต่ละคนได้ผลลัพธ์ที่คลาดเคลื่อนไปจากความเป็นจริง ดังนั้นการที่ไม่ได้เก็บข้อมูลดังกล่าวทำให้ไม่ทราบแน่ชัดว่าการทำตามเกณฑ์การคัดแยก และจัดลำดับการให้บริการผู้ป่วยฉุกเฉินตามเกณฑ์ มีประสิทธิภาพหรือไม่ ครอบคลุมการเจ็บป่วยฉุกเฉินในปัจจุบันหรือไม่ และทำให้ไม่สามารถวิเคราะห์ปัญหาแฝงที่เกิดขึ้นได้ เช่น จุดที่เกิดเหตุมากที่สุด ณ ภูมิภาคของอุบัติเหตุ ก็สามารถเป็นจุดที่ต้องให้เจ้าหน้าที่ตำรวจสามารถเฝ้าระวังได้ หรือภาคไหนที่มีอาการของโรคอะไรมากที่สุด เพื่อที่จะเตรียมสถานพยาบาลรองรับ หรือให้ความรู้เกี่ยวกับโรคนั้น ๆ ได้ ทำให้ลดความเสี่ยงต่าง ๆ ลงได้ เพื่อเป็นการเพิ่มประสิทธิภาพ เพิ่มความเร็ว และจัดเก็บข้อมูลเพื่อทำไปวิเคราะห์ต่อไป ผู้วิจัยจึงพัฒนาต้นแบบระบบสนับสนุนการตัดสินใจเกณฑ์การคัดแยก และจัดลำดับการให้บริการผู้ป่วยฉุกเฉินตามเกณฑ์ กรณีศึกษาของสถาบันการแพทย์ฉุกเฉินแห่งชาติ ในการใช้งานที่ศูนย์หลักก่อน แล้วจึงขยายไปยังศูนย์ต่าง ๆ ต่อไป

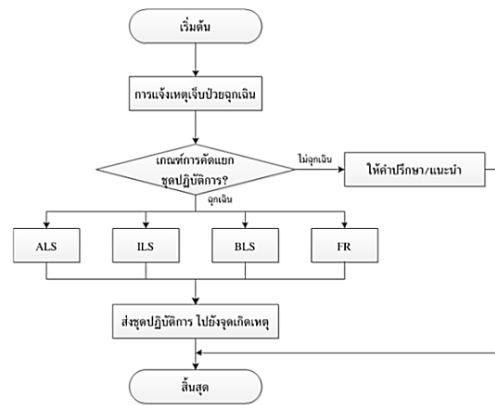
2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในการพัฒนาต้นแบบระบบนี้ ผู้วิจัยได้ศึกษาทฤษฎีในหัวข้อต่าง ๆ ที่เกี่ยวข้องดังต่อไปนี้

2.1 ระบบเดิม

ผู้วิจัยได้ทำการศึกษาภาพรวมของกระบวนการรับแจ้งเหตุ

[1] แสดงขั้นตอนดังภาพที่ 1



ภาพที่ 1: ภาพรวมของกระบวนการรับแจ้งเหตุ

จากภาพที่ 1 เป็นการแสดงขั้นตอนการรับแจ้งเหตุ โดยเริ่มตั้งแต่เจ้าหน้าที่รับแจ้งเหตุทางโทรศัพท์ และประเมินสถานการณ์ตามระดับความรุนแรงของอาการ ตามเกณฑ์การคัดแยก และจัดลำดับการให้บริการผู้ป่วยฉุกเฉินตามเกณฑ์ จนกระทั่งทราบว่าต้องส่งชุดปฏิบัติการชุดไหน ไปยังสถานที่ที่แจ้งเหตุ เพื่อดูแลรักษาผู้ป่วยฉุกเฉินเบื้องต้นก่อนถึงสถานพยาบาล

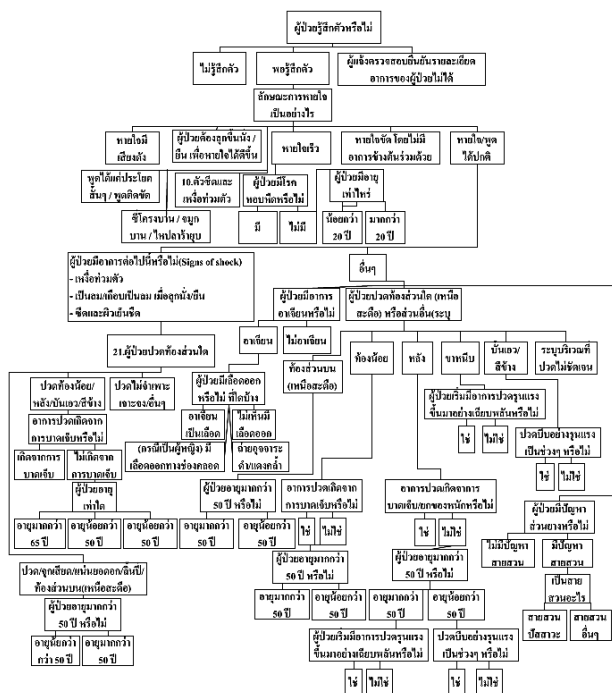
2.2 เกณฑ์การคัดแยก และจัดลำดับการให้บริการผู้ป่วย

ฉุกเฉินตามเกณฑ์

ปวดท้อง/ท้อ/เจ็บร้าว/ ขาหนีบ	เหนื่อย/อาเจียน/แพ้คลื่น คอ/แขนขา/เนื้อเยื่อ/ ปฏิกิริยาภูมิแพ้	สัตว์กัด	เลือดออก(ไม่มีสาเหตุ)
หายใจลำบาก/คิดขัด	หัวใจหยุดเต้น	เจ็บแน่นทรวงอก/หัวใจ/ มีปัญหาด้านหัวใจ	สาธิต/จุดกึ่งทางเดิน หายใจ
เมา/อาเจียน	ภาวะฉุกเฉินเหตุ สิ่งแวดล้อม	ปวดศีรษะ/ภาวะกักปัสสาวะ/ ทางคาบ/ จมูก คอ	คลื่นไส้/ภาวะทางจิต ประสาท/อารมณ์
พิษ รับประทานขนาด	มีครรภ์/คลอด/บริเวณ	ชัก/มีสัญญาณบอก เหตุการชัก	ปวด/อ่อนเพลีย/อัมพาต เรื่องอื่น ๆ
อัมพาต(กล้ามเนื้ออ่อน แรง/เฉื่อยท้อ)	ไม่รู้สติ/ไม่ตอบสนอง/ หมดสติ/จุก	เด็ก	ถูกทำร้าย
ไหม้/ลวก/เหตุจากความ ร้อน/สารเคมี/ไฟฟ้าช็อต	คนนำเจ้าหน้าที่คิดว่า จมน้ำ/บาดเจ็บเหตุค้ำ/ท่า งน้ำ	อุบัติเหตุกลุ่ม/อุบัติเหตุ/ เจ็บปวด	อุบัติเหตุอาชญากรรม

ภาพที่ 2: กลุ่มอาการทั้งหมด

ผู้วิจัยได้ทำการศึกษา พบว่ามีการจัดทำกลุ่มอาการไว้ทั้งสิ้น 24 อาการ ดังภาพที่ 2 โดยแต่ละกลุ่มอาการจะมีเกณฑ์การคัดแยก และจัดลำดับการให้บริการผู้ป่วยฉุกเฉินตามเกณฑ์ที่แตกต่างกัน เช่น กลุ่มอาการดังกล่าวอาการปวดท้อง/หลัง/เชิงกราน/ขาหนีบ แสดงดังภาพที่ 3



ภาพที่ 3: เกณฑ์การคัดแยก และจัดลำดับการให้บริการผู้ป่วยฉุกเฉิน

ตามเกณฑ์ กลุ่มของอาการปวดท้อง/หลัง/เชิงกราน/ขาหนีบ

ภาพที่ 3 เป็นการแสดงอาการปวดท้อง/หลัง/เชิงกราน/ขาหนีบ จะเห็นได้ว่า มีเกณฑ์การคัดแยก และจัดลำดับการให้บริการผู้ป่วยฉุกเฉินตามเกณฑ์ที่ซับซ้อน และใช้เวลานานในการวิเคราะห์ ระบบปัจจุบันที่ใช้งานอยู่มีความยุ่งยากในการปฏิบัติงานจริง ไม่ทันต่อเหตุการณ์ จึงทำให้ผู้ป่วยปฏิบัติงานใช้ประสิทธิภาพของตนเองในการวิเคราะห์ ซึ่งอาจทำให้ผลการวิเคราะห์ของผู้ปฏิบัติงานแต่ละคนได้ผลลัพธ์ที่คลาดเคลื่อนไปจากความเป็นจริง และผู้ปฏิบัติงานต้องกลับมาบันทึกข้อมูลบนระบบเพื่อออกรายงานตามหลัง ซึ่งการปฏิบัติงานเช่นนี้ไม่ถูกต้องตามลำดับการทำงาน ดังนั้น ระบบใหม่ จะทำการวิเคราะห์ข้อมูล และออกรายงานเพื่อให้ชุดปฏิบัติการเข้าไปยังจุดเกิดเหตุ

2.3 ศึกษางานวิจัยที่เกี่ยวข้อง

2.3.1 ออนโทโลยี

Noy และ Mcguinness [2] ได้นำเสนอวิธีการสร้างออนโทโลยี ซึ่งมี 7 ขั้นตอน ดังนี้ การกำหนด Domain และขอบเขตของออนโทโลยี การพิจารณาออนโทโลยีที่มีอยู่แล้ว การกำหนดคำสำคัญ การกำหนด Class และลำดับชั้นของ Class การกำหนด properties ของ classes—slots การกำหนด facets ของ slots และการสร้าง Instances

ปาทีตดา และคณะ [3] ได้พัฒนาระบบสืบค้นข้อมูลการท่องเที่ยวโดยใช้เทคโนโลยีออนโทโลยี และ Sematic Web โดยมีวัตถุประสงค์ในการศึกษาและจำแนกข้อมูลที่เกี่ยวข้องกับการท่องเที่ยว จากนั้นได้ทำการสร้างแบบจำลองออนโทโลยีข้อมูลการท่องเที่ยว โดยประกอบด้วย Class ต่างๆ และใช้ Sematic Web เป็นเครื่องมือในการอธิบายผลการสืบค้น

ชาลินี [4] ทำการสร้างต้นแบบออนโทโลยีพันธุ์กล้วยไม้สกุลหวาย โดยสร้างโมเดลความรู้และนิยามแนวคิดเกี่ยวกับกล้วยไม้สกุลหวายในแบบคลาส ซึ่งมี RDF (Resource Description Framework) Schema (RDFS) เป็นมาตรฐานที่ใช้ในการนิยามหรือกำหนดโครงสร้างเป็นลำดับขั้น และกำหนดความสัมพันธ์ระหว่างคลาส โดยใช้โปรแกรมโปรทีเจ (Protégé) ในการสร้างองค์ความรู้

วิชุดา และคณะ [5] ได้ทำการศึกษาปัจจัยที่เกี่ยวข้องปัจจัยที่มีผลกระทบต่อการวิเคราะห์ข่าวออนไลน์ภาษาไทย พิจารณาคำสำคัญซึ่งมีผลต่อการวิเคราะห์เนื้อหาข่าวออนไลน์ภาษาไทย โดยการหาค่าความถี่ของคำที่ปรากฏในเนื้อหาของข่าว และทำการพัฒนาฐานความรู้ออนโทโลยีสำหรับอธิบายสิ่งต่างๆ ที่ถือเป็นองค์ประกอบสำคัญในเนื้อหาข่าว ส่งผลให้ระบบสามารถวิเคราะห์เนื้อหา และอนุมานข้อมูลต่างๆ ที่เกี่ยวข้องในเนื้อหาของข่าวออนไลน์ภาษาไทยได้อย่างมีประสิทธิภาพ และทำการทดสอบฐานความรู้ออนโทโลยีโดยผู้เชี่ยวชาญ ซึ่งมีผลทดสอบอยู่ในระดับดี สามารถสรุปได้ว่าฐานความรู้ออนโทโลยีที่พัฒนาขึ้นมีการออกแบบที่ถูกต้องและสอดคล้องสำหรับนำไปประยุกต์ใช้วิเคราะห์ข่าวออนไลน์ภาษาไทยโดยอัตโนมัติได้อย่างเหมาะสม

สุภัทรา [6] พัฒนาระบบสนับสนุนการวิเคราะห์ปัญหาสำหรับเจ้าหน้าที่ Helpdesk โดยเทคโนโลยีออนโทโลยี ซึ่งองค์ความรู้ที่สร้างขึ้นสามารถกำหนดปัญหา ระบุปัญหาวิเคราะห์ปัญหา และระบุขั้นตอนการแก้ไขปัญหาที่เกิดขึ้นเกี่ยวกับเครื่องคอมพิวเตอร์แบบตั้งโต๊ะและคอมพิวเตอร์แบบพกพาได้ โดยการสร้างความองค์ความรู้ และนิยามแนวคิดเกี่ยวกับปัญหาของเครื่องคอมพิวเตอร์ในแบบคลาส ซึ่งมี RDF เป็นมาตรฐานที่ใช้ในการนิยาม หรือกำหนดโครงสร้างเป็นลำดับขั้น บรรยายลักษณะของข้อมูลในรูปแบบคลาสโดยใช้

โปรแกรมโปรทีเจ (Protégé) ทำการสร้างองค์ความรู้ การประเมินโดยภาพรวมของผู้เชี่ยวชาญ และผู้ใช้งานระบบสนับสนุนการวิเคราะห์ปัญหาสำหรับเจ้าหน้าที่ Helpdesk โดยเทคโนโลยีออนโทโลยี มีคุณภาพและความพึงพอใจในระดับดี สามารถนำระบบไปใช้งานเพื่อวิเคราะห์และสืบค้นข้อมูลได้

Borth และคณะ [7] ได้พัฒนาระบบ SentiBank โดยใช้ออนโทโลยีการนำเสนอข้อมูลขนาดใหญ่ สำหรับแยกแยะจำแนกบนเครือข่ายสังคมออนไลน์ เริ่มจากคลังแห่งอารมณ์ของ Plutchik ซึ่งประกอบด้วย 24 อารมณ์เพื่อค้นหา tags ที่เกี่ยวข้อง ไม่ว่าจะเป็นรูปภาพหรือวิดีโอ โดยให้เครื่องเรียนรู้คำนาม และกริยาที่เชื่อมต่อกัน ในการคัดแยกจะใช้ค่า F-Score มากกว่า 0.6 จากนั้นก็แสดงผลโดยใช้ Treemap สำหรับการนำไปสู่กลุ่มของรูปภาพผ่านทางออนโทโลยี Treemap ได้นำเสนอบทสรุปของโหนดในแต่ละระดับของออนโทโลยี โดยสี และขนาดจะแสดงถึงคุณสมบัติเพิ่มเติม เช่น อารมณ์ และปริมาณข้อมูล เป็นต้น

2.3.2 Tree Map

Ziegler และคณะ [8] ได้ทำการรวบรวมความคิดเห็นของลูกค้านำเป็นข้อมูลแบบไม่มีโครงสร้าง โดยทำระบบต้นแบบที่ช่วยในการวิเคราะห์กลุ่มในรูปแบบของกราฟิก การแสดงผลการวิเคราะห์อยู่บนพื้นฐานของ Treemap และสามารถคลิกลงไปดูรายละเอียดในหัวข้อที่สนใจได้อย่างรวดเร็ว

Atterer และคณะ [9] ได้นำเสนอแนวคิดในการมองเห็นความสำคัญของแต่ละบทความในบล็อกได้อย่างรวดเร็วเพียงครั้งเดียว โดยปกติพื้นที่การนำของเว็บไซต์ที่มีการจัดเป็นเมนูลำดับชั้นของหน้าหลักและหน้าย่อย ๆ งานวิจัยนี้ได้นำเสนอแนวคิดใหม่โดยพิมพ์แบบโต้ตอบของพื้นที่นำทางซึ่งตอบสนองความต้องการพิเศษของเว็บไซต์ที่มีลำดับชั้นแบบที่เชื่อมโยงไปยังเว็บไซต์ที่เกี่ยวข้องโดยซุ่มเข้าไปในบทความในบล็อกอื่น ๆ ที่เชื่อมโยงไปยังบทความบล็อกปัจจุบัน หรือเชื่อมโยงไปยังบทความที่อยู่ลึก ๆ เหล่านั้นได้

Xu และคณะ [10] ได้เสนอแนวคิดในการใช้ข้อมูลที่เก็บบนฐานข้อมูล จากนั้นได้นำ Treemap มาใช้ในการนำเสนอและวิเคราะห์กลุ่มของไฟล์ดิจิทัลส่วนตัวที่ถูกสะสมไว้ ซึ่งได้ทำการพัฒนาและทดสอบโดยใช้เวลา 14 ปีในองค์กรขนาด

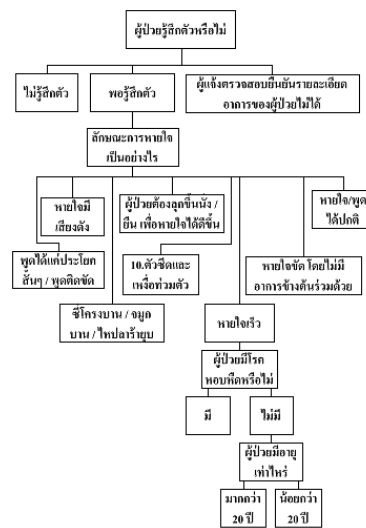
กลางแห่งหนึ่ง เพื่อดูว่าพนักงานในองค์กรจัดเก็บเอกสารดิจิทัลที่อยู่ในไดเรกทอรีที่ใช้ร่วมกันบนเครื่องเซิร์ฟเวอร์ที่เชื่อมต่อกันด้วยเครือข่ายขององค์กร พบว่าไดเรกทอรีที่ใช้ร่วมกันมีความซับซ้อน แบ่งได้เป็น 28 ลำดับ มีจำนวน 1,772 ไดเรกทอรีย่อย ๆ (sub-directories) โดยเริ่มจัดเก็บข้อมูลโครงสร้างของไดเรกทอรีไว้บนฐานข้อมูล แล้วจึงนำไปใช้กับการแสดงผลแบบ Treemap ความแตกต่างของเอกสารของพนักงานที่ถูกเก็บจะถูกแสดงผลเป็นสีส้มต่าง ๆ ซึ่งสามารถทำให้วิเคราะห์กลุ่มของพนักงานที่ใช้งานในช่วงระยะเวลาต่าง ๆ ได้ดีมาก

3. วิธีการดำเนินการ

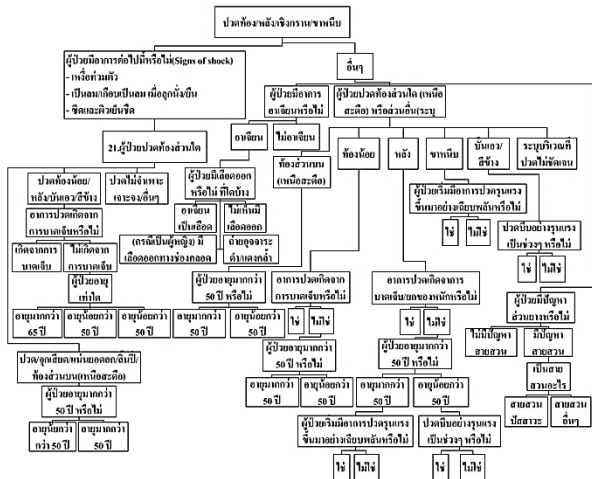
ผู้วิจัยได้ออกแบบวิธีการดำเนินการ โดยเริ่มจากนำเกณฑ์การคัดแยก และจัดลำดับการให้บริการผู้ป่วยฉุกเฉินตามเกณฑ์มาออกแบบโดยใช้ออนโทโลยี จากนั้นทำการส่งออกข้อมูลออกมาเป็น XML ไฟล์ และนำมาแสดงผลด้วยวิธีการ Treemap ดังนี้

3.1 การออกแบบเกณฑ์การคัดแยก และจัดลำดับการให้บริการผู้ป่วยฉุกเฉินตามเกณฑ์

จากการศึกษาเกณฑ์การคัดแยก และจัดลำดับการให้บริการผู้ป่วยฉุกเฉินตามเกณฑ์ของทุกกลุ่มอาการแล้ว พบว่าจะมีเกณฑ์การคัดแยก และจัดลำดับการให้บริการผู้ป่วยฉุกเฉินตามเกณฑ์มี 2 ส่วน คือส่วนที่เป็นพื้นฐาน และส่วนที่เป็นตามกลุ่มอาการนั้นๆ ดังภาพที่ 4 และภาพที่ 5



ภาพที่ 4: เกณฑ์การคัดแยก และจัดลำดับการให้บริการผู้ป่วยฉุกเฉินตามเกณฑ์ ส่วนที่เป็นพื้นฐาน



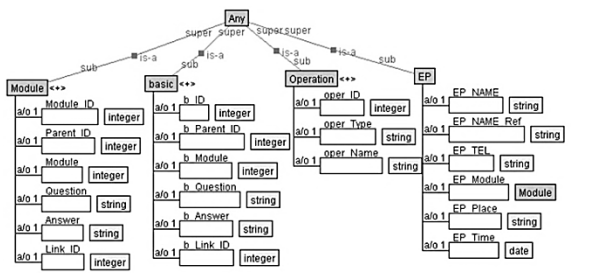
ภาพที่ 5: เกณฑ์การคัดแยก และจัดลำดับการให้บริการผู้ป่วยฉุกเฉินตามเกณฑ์ ส่วนที่ตามกลุ่มอาการนั้น ๆ

3.2 การออกแบบฐานความรู้ โดยใช้ออนโทโลยี

การออกแบบฐานความรู้ มีขั้นตอนดังนี้

- 3.2.1 ทำการกำหนดจุดประสงค์ของออนโทโลยี
- 3.2.2 พิจารณาออนโทโลยีกลับมาใช้ใหม่
- 3.2.3 กำหนดคำสำคัญ
- 3.2.4 กำหนดคลาส และซับคลาส
- 3.2.5 กำหนดคุณสมบัติของคลาส
- 3.2.6 กำหนดข้อจำกัดของคลาส
- 3.2.7 สร้างตัวอย่าง

ขั้นตอนการออกแบบฐานความรู้ โดยใช้ออนโทโลยีแสดงดังภาพที่ 6 เมื่อออกแบบแล้วจะได้ Class ต่าง ๆ ดังตารางที่ 1



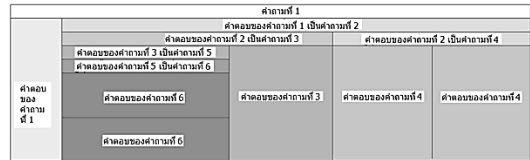
ภาพที่ 6: การออกแบบฐานความรู้ โดยใช้ออนโทโลยี

ตารางที่ 1: แสดง Class ที่ออกแบบ

ลำดับ	ชื่อคลาส	ความหมาย
1	EP (Emergency Patient)	ข้อมูลของผู้แจ้งเหตุ
2	Module	ข้อมูลของเกณฑ์
3	Basic	ข้อมูลของเกณฑ์ พื้นฐาน
4	Operation	ข้อมูลของชุดปฏิบัติการ

3.3 การออกแบบการแสดงผลการคัดแยก และจัดลำดับการให้บริการผู้ป่วยฉุกเฉินตามเกณฑ์ โดยใช้ Tree Map

ผู้วิจัยออกแบบการแสดงผลของเกณฑ์การคัดแยก และจัดลำดับการให้บริการผู้ป่วยฉุกเฉินตามเกณฑ์ โดยใช้ Tree Map ดังภาพที่ 7



ภาพที่ 7: ออกแบบหน้าจอการแสดงผล

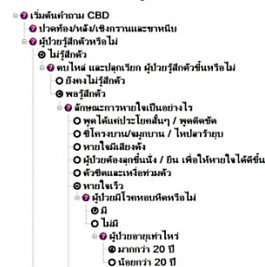
ภาพที่ 7 เป็นการแสดงผลการออกแบบหน้าจอการแสดงผล โดยคำถามหลักจะเป็นหัวตาราง และถ้าคำตอบของคำถามหลักไม่มีคำถามต่อไป จะแสดงคำตอบและชุดปฏิบัติการ แต่ถ้ามีคำถามต่อไป จะแสดงเป็นหัวตารางย่อยๆ เรียงต่อกัน

3.4 การพัฒนาระบบ

ทำการพัฒนาระบบฯ ในรูปแบบ Web Application และใช้ภาษา PHP, CSS และ JQuery ในการพัฒนา

4. ผลการดำเนินการวิจัย

ระบบเดิมนั้นเป็นการแสดงผลคำถามที่ของเกณฑ์การคัดแยก และจัดลำดับการให้บริการผู้ป่วยฉุกเฉินตามเกณฑ์แบบเป็นลำดับขั้น โดยผู้รับแจ้งเหตุต้องคลิกที่ละคำถาม และเลือกคำตอบไปที่ละคำถาม จึงจะได้ชุดปฏิบัติการ ดังภาพที่ 8 ซึ่งเป็นตัวอย่างการแสดงรูปแบบคำถามของระบบเดิม ซึ่งจากภาพถ้าผู้แจ้งเหตุแจ้งอาการว่า หายใจเร็ว -> มีโรคหอบหืด -> อายุมากกว่า 20 ปี จะต้องผ่านการคลิก หรือเลือกตอบคำถามทั้งหมด 5 ครั้ง หรือในกรณีอื่นๆ อาจจะต้องมีการคลิกมากกว่า 5 ครั้ง จึงจะได้ชุดปฏิบัติการเพื่อส่งไปยังจุดเกิดเหตุ แต่การแสดงผลในรูปแบบของ Tree Map นั้นสามารถลดขั้นตอน หรือแสดงคำถามให้อยู่ในหน้าจอเดียวกันได้



ภาพที่ 8: เป็นการแสดงรูปแบบคำถามของระบบเดิม

ส่วนหน้าจอที่ได้จากการพัฒนาระบบใหม่นั้น แสดงดังภาพที่ 9 จะเห็นได้ว่าไม่ต้องคลิกคำถามลงไปหลายระดับเพื่อให้ได้ชุดปฏิบัติการ ดังนั้นเมื่อเปรียบเทียบภาพที่ 8 แล้วจะพบว่าการปฏิบัติงานลดลงจาก 5 ครั้งเหลือ 1 ครั้ง ทำให้ผู้ปฏิบัติงานใช้เวลาในการคัดกรองผู้ป่วยฉุกเฉินได้เร็วขึ้น

ผู้บริหารส่วนรวม				
พจนานุกรม				
ลักษณะการดำเนินงานเป็นอย่างไร				
ผู้บริหารส่วนรวม	พจนานุกรม / พจนานุกรม (P 1.1.1.2)	พจนานุกรม / พจนานุกรม (P 1.1.1.2)	ตัวชี้วัด / ตัวชี้วัด (P 1.1.1.2)	พจนานุกรม / พจนานุกรม (P 1.1.1.2)
ผู้บริหารส่วนรวม	พจนานุกรม / พจนานุกรม (P 1.1.1.2)	พจนานุกรม / พจนานุกรม (P 1.1.1.2)	พจนานุกรม / พจนานุกรม (P 1.1.1.2)	พจนานุกรม / พจนานุกรม (P 1.1.1.2)
ผู้บริหารส่วนรวม	พจนานุกรม / พจนานุกรม (P 1.1.1.2)	พจนานุกรม / พจนานุกรม (P 1.1.1.2)	พจนานุกรม / พจนานุกรม (P 1.1.1.2)	พจนานุกรม / พจนานุกรม (P 1.1.1.2)
ผู้บริหารส่วนรวม	พจนานุกรม / พจนานุกรม (P 1.1.1.2)	พจนานุกรม / พจนานุกรม (P 1.1.1.2)	พจนานุกรม / พจนานุกรม (P 1.1.1.2)	พจนานุกรม / พจนานุกรม (P 1.1.1.2)

ภาพที่ 9: เป็นการแสดงรูปแบบคำถามของระบบใหม่

5. สรุปผลการวิจัย และข้อเสนอแนะ

ผู้วิจัยได้พัฒนาระบบต้นแบบระบบสนับสนุนการตัดสินใจเกณฑ์การคัดแยก และจัดลำดับการให้บริการ ผู้ป่วยฉุกเฉินตามเกณฑ์ กรณีศึกษาของสถาบันการแพทย์ฉุกเฉินแห่งชาติ ซึ่งช่วยลดขั้นตอน ลดความคลาดเคลื่อนในการวิเคราะห์โดยใช้ประสบการณ์ของแต่ละบุคคล เพิ่มความรวดเร็ว ง่ายต่อการใช้งาน และมีมาตรฐานในการตัดสินใจเดียวกัน

ข้อเสนอแนะในอนาคต เมื่อทำการรวบรวมข้อมูลเป็นระยะเวลา 1 ปีหรือมากกว่าจะทำให้ทราบถึงแนวโน้มการเกิดเหตุ แนวโน้มของอาการต่าง ๆ ที่เกิดขึ้นมากที่สุด และทำให้เตรียมพร้อมที่จะรับมือ หรือสามารถเพิ่มศูนย์สำหรับช่วยเหลือให้เพียงพอกับความต้องการ อย่างไรก็ตามระบบใหม่นี้ต้องใช้เวลาในการยอมรับจากเจ้าหน้าที่ การทดสอบกับสถานการณ์จริง และการฝึกอบรม ซึ่งต้องใช้เวลามากกว่า 1 ปี ระบบจึงสามารถนำไปใช้งานจริงได้ เพราะมีความละเอียดอ่อนและข้อบังคับต่าง ๆ ด้วย

เอกสารอ้างอิง

[1] สถาบันการแพทย์ฉุกเฉินแห่งชาติร่วมกับแผนงานร่วมสร้างเสริมสุขภาพทั้งระบบหลักประกันสุขภาพถ้วนหน้า (พรส.). เพื่อสำนักงานกองทุนสนับสนุนการสร้างเสริมสุขภาพ (สสส.), *คู่มือการจัดการและปฏิบัติงานในระบบรับแจ้งเหตุผู้ป่วยฉุกเฉิน ในระบบการแพทย์ฉุกเฉิน พ.ศ. 2551*. กรุงเทพฯ . 2552

[2] N. F. Noy and D. L. McGuinness, "Ontology Development 101: A Guide to Creating Your First Ontology", Stanford

Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880, March 2001.

[3] ปาติตดา สุขสมบุญรณ์ การ์เซีย, อัจฉรา หลีระพงส์ และนันทิยา อริยะพิชัย, *การประยุกต์ใช้เทคโนโลยีออนไลน์ และ Semantic Web สำหรับระบบสืบค้นสารสนเทศการท่องเที่ยว*. มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตตรัง, 2553.

[4] ชลินี สุกรณ์, *การสร้างต้นแบบออนโทโลยีพื้นฐานด้วยไม้อัลกอริทึม*. ปัญหาพิเศษ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2554.

[5] วิชชุดา โชติรัตน์, ผุสดี บุญรอด, และศศิมาจ ฃ วิเชียร, "การพัฒนาฐานความรู้ออนโทโลยีสำหรับวิเคราะห์ข่าวออนไลน์โดยอัตโนมัติ," *วารสารเทคโนโลยีสารสนเทศ*. ปีที่ 7 (ก.ค.ช.ค. 54) : หน้า 13-18, 2554.

[6] สุกัทัพรหมสาขาณ สกลนคร, *ระบบสนับสนุนการวิเคราะห์ปัญหาสำหรับเจ้าหน้าที่ Helpdesk โดยเทคโนโลยีออนไลน์*. ปัญหาพิเศษ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2555.

[7] D. Borth et al., "SentiBank: Large-Scale Ontology and Classifiers for Detecting Sentiment and Emotions in Visual Content", In Proceedings of the 21st ACM International Conference on Multimedia (MM'13), pp.459-460., USA: ACM, 2013.

[8] C.-N. Ziegler, M. Skubaczy and M. Viermetz, "Mining and Exploring Unstructured Customer Feedback Data Using Language Models and Treemap Visualizations", In Proceedings of the 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, WI-IAT'08. USA, Vol. 01, pp. 932-937, 2008.

[9] R. Atterer and M. Tafelmayer, "Treemap-based Website Navigation for Non-hierarchical, Interlinked Sites: The Trackback Map", In Proceedings of the 5th Nordic Conference on Human-computer Interaction: Building Bridges, pp. 411-414, USA, 2008.

[10] W. Xu, M. Esteva and S. D. Jain, "Visualizing Personal Digital Collections", In Proceedings of the 10th Annual Joint Conference on Digital Libraries, JCDL '10. New York, NY, USA: ACM, pp.169-172, 2010.

ปัจจัยที่มีอิทธิพลต่อการผสมผสานการเรียนรู้แบบกำกับตนเองและการเรียนรู้ร่วมกัน

สำหรับการฝึกอบรมผู้ใช้โปรแกรม

An Analysis of Factors Influencing Hybrid Self-Regulated and Collaborative Learning for End-User Training

สิริกอร์ กรมโพธิ์ (Sirikorn Krompho)¹ มณเฑียร รัตนศิริวงศ์วุฒิ (Montean Rattanasiriwongwut)²

และนลินภัทร์ พรวัฒน์ปรียกอร์ (Nalinpat Porrawatpreyakorn)³

^{1,3}ภาควิชาเทคโนโลยีสารสนเทศ และ ²ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

sirikorn.k@nrru.ac.th, montean@kmutnb.ac.th, nalinpatp@kmutnb.ac.th

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อวิเคราะห์ปัจจัยที่มีอิทธิพลต่อการผสมผสานการเรียนรู้แบบกำกับตนเองและการเรียนรู้ร่วมกันสำหรับการฝึกอบรมผู้ใช้โปรแกรม และตรวจสอบความสอดคล้องของโมเดลกับข้อมูลเชิงประจักษ์ ปัจจัยที่ศึกษาประกอบด้วย 3 กลุ่ม ดังนี้ 1) การวิเคราะห์ความต้องการในการฝึกอบรม ประกอบด้วยคุณลักษณะของผู้รับการฝึกอบรมและลักษณะของงาน 2) ความแตกต่างระหว่างบุคคล ประกอบด้วยการรับรู้ความสามารถของตนเองในการใช้โปรแกรมประยุกต์ ความต้องการบรรลุจุดหมายการเรียนรู้ และการรับรู้ความสามารถของตนเองก่อนการฝึกอบรม และ 3) เทคนิคการเรียนรู้ โดยมุ่งเน้นกลยุทธ์การเรียนรู้แบบกำกับตนเอง เก็บรวบรวมข้อมูลโดยใช้แบบสอบถามนักศึกษาชั้นปีที่ 4 จากมหาวิทยาลัยภาครัฐ ในประเทศไทย จำนวน 641 คน เป็นเพศหญิง 480 คน เพศชาย 161 คน วิเคราะห์ข้อมูลโดยใช้โปรแกรม SPSS และ LISREL ผลการวิจัยพบว่าทุกปัจจัยมีผลทางบวกอย่างมากต่อการผสมผสานการเรียนรู้แบบกำกับตนเองและการเรียนรู้ร่วมกันสำหรับการฝึกอบรมผู้ใช้โปรแกรม ซึ่งปัจจัยเหล่านี้มีส่วนสำคัญในการออกแบบและพัฒนาการฝึกอบรมผู้ใช้โปรแกรมให้มีประสิทธิภาพมากยิ่งขึ้น

คำสำคัญ: การฝึกอบรมผู้ใช้โปรแกรม ปัจจัยที่มีอิทธิพล การเรียนรู้แบบกำกับตนเอง การเรียนรู้ร่วมกัน

Abstract

The objective of this study is to analyze factors influencing hybrid of self-regulated and collaborative learning for end-user training. Variables included in this study were Application Specific Computer Self-Efficacy (AS-CSE), needs assessment (trainee characteristics and job/task characteristics), self-regulated learning strategy, pre-training self-efficacy, and learning goal-orientation. The samples were 641 senior undergraduate students (480 women and 161 men) at public universities in Thailand. A questionnaire survey, SPSS and LISREL were used as analysis tools.

The result indicated that needs assessments, AS-CSE (spreadsheet, presentation, and word-processing), self-regulated learning strategy, learning goal-orientation, and pre-training self-efficacy have a strong positive effect on hybrid self-regulated and collaborative learning for end-user training. These factors are important to designing and developing an effective end-user training approach.

Keyword: End-User Training, Influential Factor, Self-Regulated Learning, Collaborative Learning.

1. บทนำ

ในปัจจุบันมีการใช้เทคโนโลยีสารสนเทศและระบบสารสนเทศในองค์กรมากขึ้น การฝึกอบรมเป็นวิธีการหนึ่งที่น่ามาใช้เพื่อพัฒนาบุคลากรขององค์กร [1] ในการเรียนการสอนทักษะการใช้งานคอมพิวเตอร์ให้กับผู้ใช้โปรแกรม ส่งผลให้การรับรู้ความสามารถด้านคอมพิวเตอร์และอินเทอร์เน็ตของตนเองดีขึ้น [2] การฝึกอบรมผู้ใช้โปรแกรม (End-User Training: EUT) เป็นวิธีการหนึ่งในการพัฒนาทักษะและเรียนรู้การใช้โปรแกรม เพื่อเพิ่มประสิทธิภาพของผู้ใช้คอมพิวเตอร์และเทคโนโลยีสารสนเทศ แต่ปัจจุบันยังพบปัญหาที่ส่งผลกระทบต่อกระบวนการเรียนรู้และฝึกอบรมผู้ใช้โปรแกรม เช่น กระบวนการใช้งานโปรแกรมหรือระบบสารสนเทศจากผู้โปรแกรม [3] เนื่องจากองค์กรต่าง ๆ มีการใช้ประโยชน์จากระบบสารสนเทศที่แตกต่างกัน ทั้งนี้การออกแบบและพัฒนาวิธีการฝึกอบรมผู้ใช้โปรแกรมขึ้นอยู่กับปัจจัยหลายประการด้วยกัน ปัจจัยที่สำคัญ เช่น องค์กร ความแตกต่างระหว่างบุคคล วิธีการฝึกอบรม เทคนิคการเรียนรู้ กระบวนการเรียนรู้และการปฏิสัมพันธ์ ผลการเรียนรู้โดยตรงและผลการเรียนรู้ระยะยาว [4] เป็นต้น นอกจากนี้กลยุทธ์การฝึกอบรมและการเรียนรู้อย่างมีความสำคัญต่อกระบวนการฝึกอบรมผู้ใช้โปรแกรม โดยเฉพาะกลยุทธ์การเรียนรู้แบบกำกับตนเอง (Self-Regulated Learning Strategy) ซึ่งผู้เรียนใช้เพื่อจัดการการเรียนรู้ของตนเอง [5] และการเรียนรู้ร่วมกัน (Collaborative Learning) เพื่อสนับสนุนพฤติกรรมแบ่งปันความรู้และการมีส่วนร่วมของผู้รับการฝึกอบรม [6] อย่างไรก็ตามยังคงมีหลายปัจจัยที่เกี่ยวข้องกับความแตกต่างระหว่างบุคคลซึ่งเป็นเรื่องที่ยากและท้าทายในการพัฒนาการฝึกอบรมผู้ใช้โปรแกรมที่มีประสิทธิภาพ จึงจำเป็นต้องวิเคราะห์ปัจจัยที่มีอิทธิพลต่อการผสมผสานการเรียนรู้แบบกำกับตนเองและการเรียนรู้ร่วมกันสำหรับการฝึกอบรมผู้ใช้โปรแกรม ซึ่งจะช่วยให้สามารถเป็นแนวทางในการออกแบบและพัฒนาวิธีการฝึกอบรมผู้ใช้โปรแกรมให้มีประสิทธิภาพมากยิ่งขึ้น เช่น สถาบันการศึกษาที่มุ่งเน้นให้บัณฑิตมีคุณภาพครอบคลุมการสื่อสารและเทคโนโลยีสารสนเทศ ตามกรอบมาตรฐานคุณวุฒิระดับอุดมศึกษา แห่งชาติ พ.ศ. 2552 [7] เป็นต้น

2. ทฤษฎีและวรรณกรรมที่เกี่ยวข้อง

2.1 ทฤษฎีพื้นฐาน

การฝึกอบรมผู้ใช้โปรแกรม หมายถึง โปรแกรมการฝึกอบรมสำหรับการใช้งานระบบสารสนเทศหรือโปรแกรมประยุกต์สำหรับผู้ใช้แตกต่างจากวิธีการฝึกอบรมแบบเดิมที่มีโครงสร้างสภาพแวดล้อมแบบห้องเรียน เป้าหมายของการพัฒนา EUT คือ การสร้างแรงจูงใจผู้ใช้ให้มีทักษะที่จำเป็นในการประยุกต์ใช้และการเรียนรู้เพื่อใช้งาน ในปัจจุบันหลายองค์กรมีการฝึกอบรมโดยใช้เทคโนโลยี (Technology-Based Training) เช่น E-Learning สำหรับการฝึกอบรมร่วมกับการเรียนรู้ร่วมกัน เพื่อให้บรรลุเป้าหมายการเรียนรู้ และลดค่าใช้จ่ายในการดำเนินการฝึกอบรม เป็นต้น ซึ่งเป็นวิธีการฝึกอบรมวิธีหนึ่งที่ใช้ในองค์กรมากขึ้น [3] นอกจากนี้ยังมีการพัฒนาวิธีการฝึกอบรมบนพื้นฐานทฤษฎีการเรียนรู้ทางปัญญาสังคม (Social Cognitive Theory) [8] และการฝึกอบรมผู้ใช้โปรแกรมโดยใช้เทคโนโลยี (Technology-Mediated Training Methods) เช่น การฝึกอบรมโดยใช้บทเรียนคอมพิวเตอร์ ทั้งนี้การพัฒนาวิธีการฝึกอบรมผู้ใช้โปรแกรมยังมุ่งเน้นเป้าหมายการฝึกอบรมและความแตกต่างระหว่างบุคคล [3]

2.2 งานวิจัยที่เกี่ยวข้อง

ปัจจัยที่มีอิทธิพลต่อการฝึกอบรมผู้ใช้โปรแกรมสามารถแบ่งออกเป็น 7 กลุ่ม ดังนี้ [4]

กลุ่มที่ 1 ปัจจัยด้านองค์กร เช่น การวิเคราะห์ความต้องการในการฝึกอบรม [9] จะช่วยให้ทราบข้อมูลที่เป็นสำหรับการออกแบบและพัฒนาวิธีการฝึกอบรมผู้ใช้โปรแกรม เพื่อกำหนดเป้าหมายการฝึกอบรมและนำเข้าสู่กระบวนการฝึกอบรมและการเรียนรู้

กลุ่มที่ 2 ปัจจัยด้านความแตกต่างระหว่างบุคคล ซึ่งมีผลต่อกระบวนการฝึกอบรมผู้ใช้โปรแกรม เช่น การรับรู้ความสามารถของตนเองก่อนการฝึกอบรม (Pre-training Self-Efficacy: SE) [10] การรับรู้ความสามารถของตนเองด้านคอมพิวเตอร์ (Computer Self-Efficacy: CSE) ประกอบด้วยการใช้งานคอมพิวเตอร์ทั่วไปและการใช้งานโปรแกรมเฉพาะ เช่น การใช้งานโปรแกรมประมวลผลคำ [11] และความต้องการบรรลุจุดหมายการเรียนรู้ซึ่งเป็นความตั้งใจของผู้เรียนที่จะพัฒนาความสามารถในระยะยาว [10] เป็นต้น

กลุ่มที่ 3 ปัจจัยด้านวิธีการฝึกอบรม เป็นการฝึกอบรมด้วยวิธีการต่าง ๆ เช่น วิธีการแบบครูผู้สอนหน้าชั้นเรียน การฝึกอบรมโดยใช้ e-Learning การฝึกอบรมแบบพี่เลี้ยง และวิธีการฝึกอบรมแบบผสมผสาน [12]

กลุ่มที่ 4 ปัจจัยด้านเทคนิคการเรียนรู้ เช่น วิธีการเขียนแบบพฤติกรรม [13] การเรียนรู้ผลของพฤติกรรม และการเรียนรู้จากการกระทำหรือปฏิสัมพันธ์กับสิ่งแวดล้อม [14] มีอิทธิพลทางบวกต่อการฝึกอบรมผู้ใช้โปรแกรม และการเรียนรู้ร่วมกันที่ส่งเสริมให้ผู้เรียนทำงานร่วมกันเพื่อบรรลุเป้าหมายร่วมกัน [3] การใช้กลยุทธ์การเรียนรู้แบบกำกับตนเองส่งผลให้การเรียนรู้เพื่อใช้ระบบสารสนเทศดีขึ้น [5]

กลุ่มที่ 5 ปัจจัยด้านกระบวนการเรียนรู้และการปฏิสัมพันธ์ เช่น การออกแบบส่วนติดต่อกับผู้ใช้เรียนและการปฏิสัมพันธ์ [15] เป็นต้น

กลุ่มที่ 6 ปัจจัยด้านผลการเรียนรู้โดยตรงเป็นปัจจัยที่มีอิทธิพลต่อผลการเรียนรู้ เช่น ประสิทธิภาพการเรียนรู้ (Learning Performance) [15] เป็นต้น

กลุ่มที่ 7 ปัจจัยด้านผลการเรียนรู้ระยะยาวเกี่ยวข้องกับการสนับสนุนหลังการฝึกอบรม เพื่อให้ผู้รับการฝึกอบรมมีระดับทักษะสูงขึ้น [16]

3. วิธีการดำเนินการวิจัย

การวิจัยนี้มีขั้นตอนและวิธีการที่ใช้ดังนี้

3.1 การกำหนดปัจจัยและสมมติฐานการวิจัย

จากปัจจัยที่มีอิทธิพลต่อการผสมผสานการเรียนรู้แบบกำกับตนเองและการเรียนรู้ร่วมกันสำหรับการฝึกอบรมผู้ใช้โปรแกรม โดยเฉพาะการรับรู้ความสามารถของตนเองด้าน

คอมพิวเตอร์มีผลต่อกระบวนการนำเข้าสู่การฝึกอบรม กระบวนการฝึกอบรมและการเรียนรู้ และผลการฝึกอบรม สามารถตั้งสมมติฐานได้ 7 ข้อ ดังนี้

สมมติฐานที่ 1 การรับรู้ความสามารถของตนเองในการใช้โปรแกรมประมวลผลคำจะส่งผลทางบวกต่อการฝึกอบรมผู้ใช้โปรแกรม (H1)

สมมติฐานที่ 2 การรับรู้ความสามารถของตนเองในการใช้โปรแกรมตารางคำนวณจะส่งผลทางบวกต่อการฝึกอบรมผู้ใช้โปรแกรม (H2)

สมมติฐานที่ 3 การรับรู้ความสามารถของตนเองในการใช้โปรแกรมนำเสนอผลงานจะส่งผลทางบวกต่อการฝึกอบรมผู้ใช้โปรแกรม (H3)

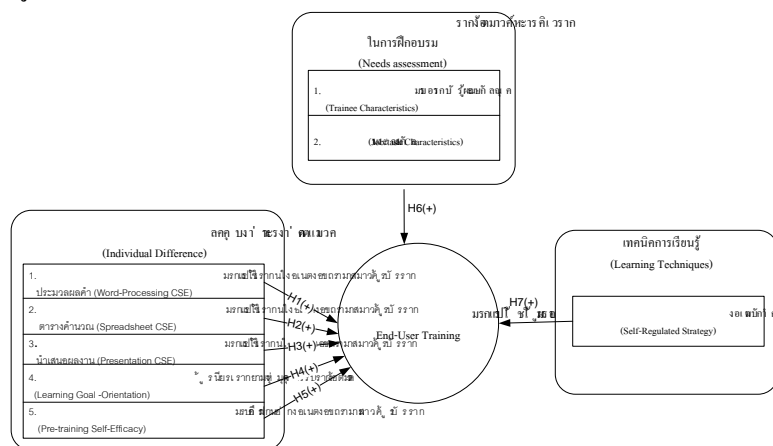
สมมติฐานที่ 4 ความต้องการบรรลุจุดหมายการเรียนรู้จะส่งผลทางบวกต่อการฝึกอบรมผู้ใช้โปรแกรม (H4)

สมมติฐานที่ 5 การรับรู้ความสามารถของตนเองจะส่งผลทางบวกต่อการฝึกอบรมผู้ใช้โปรแกรม (H5)

สมมติฐานที่ 6 การวิเคราะห์ความต้องการในการฝึกอบรมจะส่งผลทางบวกต่อการฝึกอบรมผู้ใช้โปรแกรม (H6)

สมมติฐานที่ 7 กลยุทธ์การเรียนรู้แบบกำกับตนเองจะส่งผลทางบวกต่อการฝึกอบรมผู้ใช้โปรแกรม (H7)

จากสมมติฐานข้างต้น นำมาสร้างเป็นโมเดลการวิเคราะห์ปัจจัย 3 ด้าน คือ ด้านความแตกต่างระหว่างบุคคล ด้านความต้องการในการฝึกอบรม และด้านเทคนิคการเรียนรู้ โดยกำหนดให้สมมติฐานที่ส่งผลทางบวกแทนด้วย Hx(+) โดย x แทนลำดับสมมติฐาน ดังสรุปในภาพที่ 1



ภาพที่ 1: โมเดลการวิเคราะห์ปัจจัยที่มีอิทธิพลต่อการผสมผสานการเรียนรู้แบบกำกับตนเองและการเรียนรู้ร่วมกันสำหรับการฝึกอบรมผู้ใช้โปรแกรม

3.2 การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลดำเนินการโดยใช้แบบสอบถามแบบประมาณค่า 5 ระดับ ทดลองใช้กับกลุ่มทดสอบเครื่องมือจำนวน 30 คน เพื่อทดสอบหาความเชื่อมั่น ผลการทดสอบพบว่าปัจจัยด้านความวิตกกังวลในการใช้คอมพิวเตอร์มีความเชื่อมั่นต่ำ ดังนั้นการเก็บข้อมูลจริงจึงตัดข้อคำถามในส่วนนี้ออกไป 4 ข้อ รายละเอียดแบบสอบถามปัจจัยแต่ละด้านแสดงดังตารางที่ 1

ตารางที่ 1: ค่าความเชื่อมั่น

ปัจจัย	จำนวน	α
ความวิตกกังวลในการใช้คอมพิวเตอร์	4	0.65
ความสามารถในการใช้โปรแกรมประมวลผลค่า	5	0.83
ความสามารถในการใช้โปรแกรมตารางคำนวณ	7	0.89
ความสามารถในการใช้โปรแกรมนำเสนอผลงาน	7	0.93
การวิเคราะห์ความต้องการในการฝึกอบรม	5	0.85
กลยุทธ์การเรียนรู้แบบกำกับตนเอง	9	0.89
การรับรู้ความสามารถของตนเองก่อนการฝึกอบรม	5	0.87
ความต้องการบรรลุจุดหมายการเรียนรู้	5	0.88
รวมทั้งฉบับ	47	0.97

การกำหนดประชากรและกลุ่มตัวอย่างใช้วิธีการสุ่มตัวอย่างอย่างง่ายไม่ต่ำกว่า 20 เท่าของตัวแปร จำนวน 1,000 คน กลุ่มตัวอย่างที่ใช้ในงานวิจัย คือ นักศึกษาชั้นปีที่ 4 ของมหาวิทยาลัยภาคีรัฐ ในประเทศไทย 10 แห่ง ได้แก่ มหาวิทยาลัยขอนแก่น มหาวิทยาลัยมหาสารคาม มหาวิทยาลัยเทคโนโลยีสุรนารี มหาวิทยาลัยราชภัฏเชียงใหม่ มหาวิทยาลัยราชภัฏนครราชสีมา มหาวิทยาลัยราชภัฏภูเก็ต มหาวิทยาลัยราชภัฏบุรีรัมย์ มหาวิทยาลัยราชภัฏมหาสารคาม มหาวิทยาลัยราชภัฏอุบลราชธานี และมหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน ซึ่งเก็บรวบรวมตั้งแต่วันที่ 20 สิงหาคม 2556 ถึงวันที่ 20 ตุลาคม 2556 โดยได้รับแบบสอบถามกลับมาจำนวน 641 ชุด คิดเป็นร้อยละ 64.10 จำแนกเป็นเพศหญิง 480 คน และเพศชาย 161 คน

3.3 การวิเคราะห์ข้อมูล

วิเคราะห์ค่าร้อยละ ค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน วิเคราะห์องค์ประกอบเชิงสำรวจด้วยโปรแกรม SPSS และวิเคราะห์องค์ประกอบเชิงยืนยันด้วยโปรแกรม LISREL

4. ผลการดำเนินงาน

ผลการวิเคราะห์องค์ประกอบเชิงสำรวจพบว่า เมื่อพิจารณาค่าสถิติ Bartlett's Test of Sphericity มีค่าเท่ากับ 15,551.81 ค่าดัชนีไกเซอร์-ไมเยอร์-ออลกิน (Kaiser-Meyer-Olkin: KMO) มีค่าเท่ากับ 0.96 ซึ่งมากกว่า 0.50 และเข้าสู่ 1 แสดงว่าความสัมพันธ์ระหว่างตัวแปรมีความเหมาะสมดีมากที่จะนำมาใช้ในการวิเคราะห์องค์ประกอบเชิงยืนยัน สามารถสร้างเป็นโมเดลโครงสร้างปัจจัยที่มีอิทธิพลต่อการผสมผสานการเรียนรู้แบบกำกับตนเองและการเรียนรู้ร่วมกันสำหรับการฝึกอบรมผู้ใช้โปรแกรม

ผลการวิเคราะห์องค์ประกอบเชิงยืนยันขั้นต้นที่หนึ่งพบว่า ค่าสัมประสิทธิ์การพยากรณ์ (R^2) บางตัวมีค่าน้อยกว่า 0.50 หรืออาจมีความสอดคล้องกับปัจจัยในแต่ละด้านน้อยเกินไป ผู้วิจัยจึงตัดตัวแปรออก 10 ตัวแปร สรุปได้ว่าโมเดลประกอบด้วย 7 ปัจจัย 33 ตัวแปร การตรวจสอบว่าทั้ง 7 ปัจจัย เป็นปัจจัยที่มีอิทธิพลต่อการผสมผสานการเรียนรู้แบบกำกับตนเองและการเรียนรู้ร่วมกันสำหรับการฝึกอบรมผู้ใช้โปรแกรม ผู้วิจัยได้ทำการคัดเลือกตัวแปรสังเกตได้ที่สำคัญจากแต่ละปัจจัย รายละเอียดดังตารางที่ 2

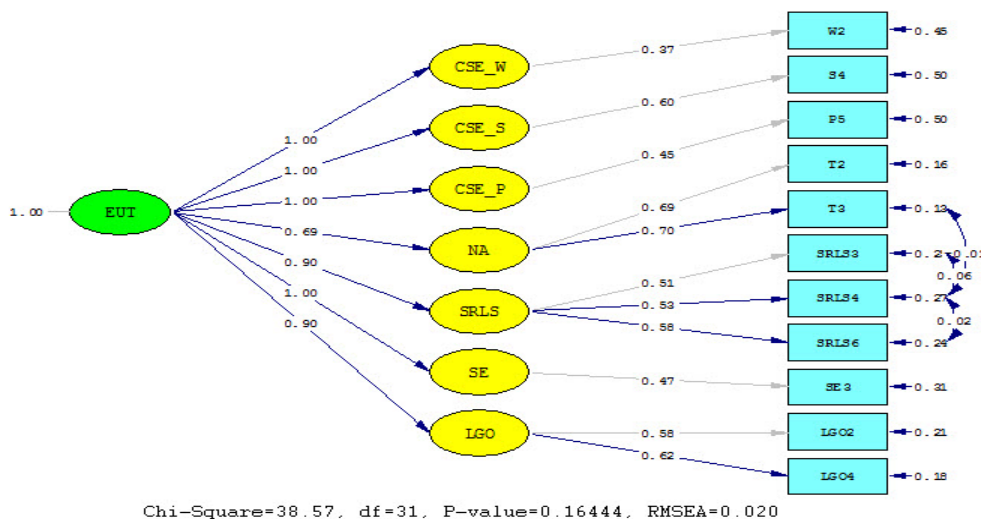
ซึ่งจากตารางที่ 2 แสดงผลการคัดเลือกตัวแปรสังเกตได้ 11 ตัวแปร จากปัจจัยที่มีอิทธิพลต่อการผสมผสานการเรียนรู้แบบกำกับตนเองและการเรียนรู้ร่วมกันสำหรับการฝึกอบรมผู้ใช้โปรแกรม 7 ปัจจัย น้ำหนักองค์ประกอบของตัวแปรทั้ง 11 ตัวแปรมีค่าเป็นบวก โดยมีค่าน้ำหนักองค์ประกอบตั้งแต่ 0.62 ถึง 0.82 และมีนัยสำคัญทางสถิติที่ระดับ 0.05 ทุกค่า แสดงว่าตัวแปรทั้งหมด 11 ตัว เป็นตัวแปรที่สำคัญ และยังสามารถพิจารณาได้จากค่าความผันแปรร่วมกับองค์ประกอบมาตรฐาน (R^2) ซึ่งมีค่ามากกว่า 0.50 สำหรับผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง พบว่า โมเดลสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ โดยพิจารณาจากค่าไค-สแควร์ มีค่าเท่ากับ 38.57 ซึ่งมีค่าความน่าจะเป็นเท่ากับ 0.16 ที่องศาอิสระเท่ากับ 31 นั่นคือ ค่าไค-สแควร์แตกต่างจากศูนย์อย่างไม่มีนัยสำคัญทางสถิติที่ระดับ 0.05 แสดงว่ายอมรับสมมติฐานหลักที่ว่า โมเดลการวัดมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์

ตารางที่ 2: ผลการวิเคราะห์องค์ประกอบเชิงยืนยันชั้นชั้นที่หนึ่ง

ตัวแปร	น้ำหนักองค์ประกอบ	R ²	สปส. คะแนนองค์ประกอบ
ปัจจัยด้านการรับรู้ความสามารถในการใช้โปรแกรมประมวลผลคำ (CSE_W)			
W2 ความสามารถในการจัดการกับย่อหน้า	0.62	0.65	0.43
ปัจจัยด้านการรับรู้ความสามารถในการใช้โปรแกรมตารางคำนวณ (CSE_S)			
S4 ความสามารถที่จะสรุปข้อมูลตัวเลขโดยใช้โปรแกรม	0.82	0.76	0.31
ปัจจัยด้านการรับรู้ความสามารถในการใช้โปรแกรมนำเสนอผลงาน (CSE_P)			
P5 ความสามารถตกแต่งพื้นหลังและสร้างแม่แบบได้	0.75	0.62	0.15
ปัจจัยด้านทฤษฎีการวิเคราะห์ความต้องการในการฝึกอบรม (NA)			
T2 ผู้รับการฝึกอบรมควรมีความกระตือรือร้นในการแสวงหาความรู้อยู่เสมอ	0.68	0.72	0.38
T3 ผู้รับการฝึกอบรมต้องมีความพร้อมทั้งทางร่างกายและจิตใจในการรับความรู้จากเครื่องมือสนับสนุนการฝึกอบรมผู้ใช้โปรแกรม	0.70	0.80	0.59
ปัจจัยด้านกลยุทธ์การเรียนรู้แบบกับตนเอง (SRLS)			
SRLS3 การตั้งคำถามกับตนเองเพื่อให้แน่ใจว่าเข้าใจสิ่งที่ได้พยายามที่จะเรียนรู้	0.63	0.58	0.28
SRLS4 ความพยายามที่จะตรวจสอบว่าสิ่งใดที่จำเป็นต้องมีการปรับปรุงมากที่สุด	0.65	0.59	0.31
SRLS6 ความพยายามที่จะผ่านในแต่ละหัวข้อและตัดสินใจสิ่งที่เรียนรู้	0.64	0.56	0.18
ปัจจัยด้านกรรับรู้ความสามารถของตนเองก่อนการฝึกอบรม (SE)			
SE3 ความสามารถใช้เครื่องมือสนับสนุนการฝึกอบรมผู้ใช้โปรแกรม ถ้าสามารถขอความช่วยเหลือเหลือจากผู้อื่นเมื่อมีปัญหา	0.63	0.74	0.64
ปัจจัยด้านความต้องการบรรลุจุดหมายการเรียนรู้ (LGO)			
LGO2 ความชอบที่จะทำงานเพื่อเรียนรู้สิ่งใหม่	0.65	0.79	0.66
LGO4 โอกาสที่จะขยายขีดความสามารถของตนเอง	0.63	0.70	0.42

โดยค่าดัชนีวัดระดับความกลมกลืน (GFI) เท่ากับ 0.99 ค่าดัชนีวัดระดับความกลมกลืนที่ปรับแก้แล้ว (AGFI) เท่ากับ

0.98 และค่าดัชนีรากของกำลังสองเฉลี่ยของเศษ (RMR) เท่ากับ 0.02 ผลการวิเคราะห์แสดงดังภาพที่ 2



ภาพที่ 2: โมเดลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับที่สอง

จากภาพที่ 2 น้ำหนักองค์ประกอบทั้ง 7 ด้าน มีค่าเป็นบวก ผลการวิเคราะห์ข้างต้นนี้ แสดงว่าผู้รับการฝึกอบรมจะ

สามารถใช้เครื่องมือสนับสนุนการฝึกอบรมผู้ใช้โปรแกรมได้ ถ้าสามารถขอความช่วยเหลือเหลือจากผู้อื่นเมื่อมีปัญหา (SE3) ควร

มีความพร้อมทั้งทางร่างกายและจิตใจในการรับความรู้จากเครื่องมือสนับสนุนการฝึกอบรมผู้ใช้โปรแกรม (T3) อีกทั้งควรมีความกระตือรือร้นในการแสวงหาความรู้อยู่เสมอ (T2) และโอกาสที่จะขยายขีดความสามารถของตนเอง (LG04) ตามลำดับ สามารถสรุปได้ว่าสมมติฐานทั้ง 7 ข้อเป็นจริง กล่าวคือองค์ประกอบทั้ง 7 ด้านนั้นส่งผลทางบวกต่อการฝึกอบรมผู้ใช้โปรแกรมจริง

5. สรุป

ปัจจัยที่ส่งผลทางบวกต่อการผสมผสานการเรียนรู้แบบกำกับตนเองและการเรียนรู้ร่วมกันสำหรับการฝึกอบรมผู้ใช้โปรแกรม ประกอบด้วย 7 ด้าน คือ ด้านการรับรู้ความสามารถในการใช้โปรแกรมประมวลผลคำ ด้านการรับรู้ความสามารถในการใช้โปรแกรมตารางคำนวณ ด้านการรับรู้ความสามารถในการใช้โปรแกรมนำเสนอผลงาน ด้านการวิเคราะห์ความต้องการในการฝึกอบรม ด้านกลยุทธ์การเรียนรู้แบบกำกับตนเอง ด้านการรับรู้ความสามารถของตนเองก่อนการฝึกอบรม และด้านความต้องการบรรลุจุดหมายการเรียนรู้ ซึ่งปัจจัยดังกล่าวมีส่วนสำคัญที่จะพัฒนาความสามารถของผู้ใช้โปรแกรมในระยะยาวเพื่อให้ได้มาซึ่งความรู้และทักษะ โดยการเรียนรู้แบบกำกับตนเองและการเรียนรู้ร่วมกัน

6. กิตติกรรมประกาศ

ขอขอบคุณทุนโครงการส่งเสริมการวิจัยในอุดมศึกษาและพัฒนามหาวิทยาลัย สำนักงานคณะกรรมการการอุดมศึกษา ที่สนับสนุนทุนการศึกษาให้แก่ สิริกร กรมโพธิ์

เอกสารอ้างอิง

- [1] R. P. Bostrom, "Technology-mediated learning: A comprehensive theoretical model," *Journal of the Association for Information Systems*, vol. 10, 2009.
- [2] G. Torkzadeh, J. C.-J. Chang, and D. Demirhan, "A contingency model of computer and Internet self-efficacy," *Information & Management*, vol. 43, pp. 541-550, 2006.
- [3] S. Gupta, R. P. Bostrom, and M. Huber, "End-user training methods: what we know, need to know," *SIGMIS Database*, vol. 41, pp. 9-39, 2010.

- [4] S. Krompho, N. Porrawatpreyakorn, S. Smachat, and M. Rattanasiriwongwut, "Factors Influencing Hybrid Self-Regulated and Collaborative Learning for End-User Training: A Systematic Literature Review," in *The 15th International Conference on Information Integration and Web-based Applications & Services (iiWAS2013)*, Vienna, Austria, 2013, pp. 189-193.
- [5] J. Gravill and D. Compeau, "Self-regulated learning strategies and software training," *Information & Management*, vol. 45, pp. 288-296, 2008.
- [6] S. Gupta, R. P. Bostrom, and R. Anson, "Do I matter?: the impact of individual differences on training process," presented at the Proceedings of the 2010 Special Interest Group on Management Information System's 48th annual conference on Computer personnel research on Computer personnel research, Vancouver, BC, Canada, 2010.
- [7] กระทรวงศึกษาธิการ, "กรอบมาตรฐานคุณวุฒิ ระดับอุดมศึกษาแห่งชาติ พ.ศ. 2552," ed, 2552.
- [8] A. Bandura, *Social foundations of thought and action: A social cognitive theory*: Prentice-Hall, Inc, 1986.
- [9] L. Olfman, R. P. Bostrom, and M. K. Sein, "A best-practice based model for information technology learning strategy formulation," presented at the Proceedings of the 2003 SIGMIS conference on Computer personnel research: Freedom in Philadelphia--leveraging differences and diversity in the IT workforce, Philadelphia, Pennsylvania, 2003.
- [10] P. D. Chatzoglou, L. Sarigiannidis, E. Vraimaki, and A. Diamantidis, "Investigating Greek employees' intention to use web-based training," *Computers & Education*, vol. 53, pp. 877-889, 2009.
- [11] O. T. Eytayo, "Do students have the relevant ICT skills they need to do their research projects," presented at the Proceedings of the 2011 conference on Information technology education, West Point, New York, USA, 2011.
- [12] N. Kaplan-Mor, C. Glezer, and M. Zviran, "A comparative analysis of end-user training methods," *Journal of Systems and Information Technology*, vol. 13, pp. 25-42, 2011.
- [13] C. C. Chen, T. Ryan, and L. Olfman, "The efficacy of behavior modeling: A comparison of face-to-face and online asynchronous software-training

methods," in *System Sciences*, 2004. *Proceedings of the 37th Annual Hawaii International Conference on*, 2004, p. 10 pp.

- [14] S. Gupta and R. P. Bostrom, "Achieving end-user training effectiveness through web-based training systems: An empirical study," Hyderabad, 2009.
- [15] W. Cheng-Hua, L. Lu-Wen, and C. Yi-Ying, "The effect of ERP software interface design and cognitive function on performance of user learning," in *Service Operations, Logistics, and Informatics (SOLI)*, 2011 *IEEE International Conference on*, 2011, pp. 225-230.
- [16] S. Snoddy Jr and D. G. Novick, "Post-training support for learning technology," in *Proceedings of the 22nd annual international conference on Design of communication: The engineering of quality documentation*, 2004, pp. 90-96.

การประยุกต์ใช้ซอฟต์แวร์ Solr เพื่อการรับแจ้งและค้นคืนปัญหา TOT iptv

On the Application of Solr software for Helpdesk and Retrieval issue TOT iptv

กฤษณ์ ศิริทรัพย์ (Krit Sirisup)¹ และณัฐวี อุตกฤษฎ์ (Nattavee Utakrit)²

^{1,2}ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

ksirisup@hotmail.com¹, nattaveeu@kmutnb.ac.th²

บทคัดย่อ

บทความนี้ได้นำเสนอ ระบบรับแจ้งและค้นคืนปัญหา TOT iptv ของ บริษัท ทีโอที จำกัด (มหาชน) โดยมีวัตถุประสงค์เพื่อพัฒนาการบันทึกข้อมูลจากรูปแบบเอกสารเป็นการบันทึกข้อมูลเข้าสู่คอมพิวเตอร์แทน ทำให้สะดวกต่อการค้นหา และสำรองข้อมูล โดยในการจัดเก็บและค้นหาได้นำทฤษฎีการค้นคืนสารสนเทศ ได้แก่ การตัดคำภาษาไทย มาช่วยทำการสกัดข้อความที่จะค้นหา เพื่อเข้าสู่ขั้นตอนการวิเคราะห์ และการทำตัวแทนเอกสารหรือดัชนี โดยวิธีดังกล่าวจะช่วยเพิ่มประสิทธิภาพในการสืบค้น ทำให้ค้นหาข้อมูลได้สะดวก รวดเร็ว และได้ข้อมูลที่ตรงกับความต้องการ หลังจากพัฒนาระบบรับแจ้งและค้นคืนปัญหา TOT iptv ได้ทำการทดลองใช้งานจริง และทำการประเมินประสิทธิภาพของระบบผลการวัดประสิทธิภาพจากการทดสอบโดยผู้ใช้ จากคำที่ใช้ในการค้นหาปัญหา TOT iptv พบว่า ประสิทธิภาพของการค้นหามีความแม่นยำ (Precision) ร้อยละ 74 และค่าการเรียกคืน (Recall) ร้อยละ 76 ดังนั้นระบบที่พัฒนาขึ้นมีประสิทธิภาพในการค้นคืนอยู่ในระดับดี

คำสำคัญ: ปัญหา TOT iptv การค้นคืนสารสนเทศ การตัดคำภาษาไทย

Abstract

This paper presents Helpdesk and Retrieval issue system for TOT iptv, a product of TOT Public Company Limited. The purpose of this paper is to develop the method of how to keeping records from the original way of writing on paper to using the computer which makes it more convenient to search for data and also to back up.

In order to collect or search data, information retrieval theory is being introduced. This theory uses Thai Word Segmentation method to extract the exact word which is needed to be searched and then sends them to an analysis process. In addition, indexing is implemented to help increase the efficiency of searching by making the searching process more conveniently, rapidly and accurately. After the system having been developed, testing by users and evaluating the efficiency of the system are performed. The result shows that the efficiency by precision of searching process is approximately 74 percent with the recall of approximately 76 percent. Therefore, this system efficiency is said to be in a fine level.

Keyword: TOT iptv issue, Information Retrieval, Thai Word Segmentation.

1. บทนำ

ปัจจุบันสารสนเทศได้เข้ามามีบทบาท และมีความสำคัญอย่างยิ่งในการประกอบกิจกรรมต่างๆ ในชีวิตประจำวัน ส่งผลให้ทุกองค์กรได้นำระบบสารสนเทศเข้ามาประยุกต์ใช้ในกระบวนการทำงาน เพื่อช่วยในการบริหารจัดการให้มีความเป็นระบบ และมีประสิทธิภาพมากยิ่งขึ้น โดยองค์กรต่างๆ ได้นำระบบสารสนเทศมาประยุกต์ใช้ในงานต่าง ๆ มากมาย อาทิ เช่น บริษัท ทู วู้ชส์ จำกัด (มหาชน) ได้นำระบบสารสนเทศมาช่วยในการควบคุมภายใน ได้แก่ ระบบการเงิน บัญชี จัดซื้อ ทำให้การบริหารจัดการเป็นไปอย่างรวดเร็ว ถูกต้อง [1] นอกจากนี้ยังมีบริษัทอีกมากมายที่ได้นำระบบสารสนเทศเข้ามาใช้ในองค์กร จึงกล่าวได้ว่า “คนเราต้องการสารสนเทศเพื่อ

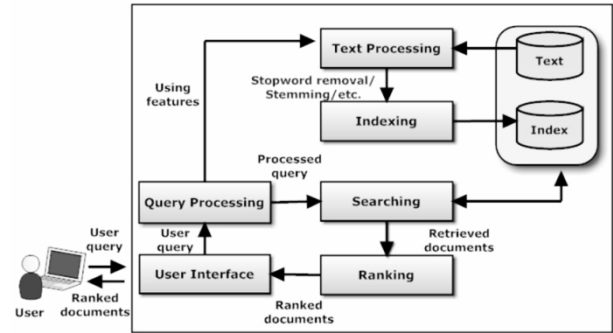
แก้ปัญหา สารสนเทศจึงเป็นกลไกสำคัญในการปฏิบัติงานให้บรรลุวัตถุประสงค์ที่กำหนดไว้” [2] ด้วยเหตุผลข้างต้น จึงสอดคล้องกับการนำสารสนเทศเข้ามาประยุกต์ใช้ในการให้บริการผลิตภัณฑ์ TOT iptv ของบริษัท ทีโอที จำกัด (มหาชน) จากลักษณะการทำงานที่เป็นอยู่ในปัจจุบัน จะเป็นการจัดเก็บข้อมูลในรูปแบบเอกสารจำนวนมาก จึงเกิดความไม่สะดวกในการจัดเก็บและสืบค้น โดยในการค้นหาข้อมูลของปัญหา TOT iptv ที่มีอยู่เดิม และที่จะเพิ่มขึ้นในอนาคต ทำให้การค้นหาข้อมูลที่มีอยู่เป็นจำนวนมาก เป็นไปอย่างยากลำบาก และต้องใช้ระยะเวลาาน รวมถึงผลลัพธ์ของการค้นหาที่ได้ อาจไม่ตรงตามความต้องการ เนื่องจากการแปลความหมายของคำที่ใช้ในการค้นหา อาจมีความเข้าใจที่แตกต่างกัน ทำให้ข้อมูลที่ได้จากการค้นหาไม่ตรงกับความต้องการของผู้สืบค้นข้อมูล

จากปัญหาดังกล่าว จึงเป็นที่มาให้ผู้วิจัยทำการพัฒนาระบบรับแจ้งและสืบค้นปัญหา TOT iptv ขึ้น ซึ่งระบบดังกล่าวพัฒนาขึ้นด้วยการใช้เทคนิคการตัดคำ และการทำตัวแทนเอกสารหรือดัชนี (Indexing) เช่นเดียวกับ Twitter ที่เป็นที่ยอมรับในโลกลงคมออนไลน์ในขณะนี้ ที่ได้นำเทคโนโลยีดังกล่าวมาใช้ในการค้นหาข้อมูล โดย Twitter สามารถรองรับการค้นหาข้อมูลที่มีมากถึง 12,000 ครั้งต่อวินาที หรือ 1 พันล้านครั้งต่อวัน [3] จากเหตุผลดังกล่าวจึงเป็นที่มาที่ทำให้ผู้วิจัยเลือกใช้เทคนิคการตัดคำ และการทำตัวแทนเอกสารหรือดัชนี มาประกอบกับการสืบค้นข้อมูล เพื่อให้การสืบค้นข้อมูลมีประสิทธิภาพมากยิ่งขึ้น

2. ทฤษฎีและวรรณกรรมที่เกี่ยวข้อง

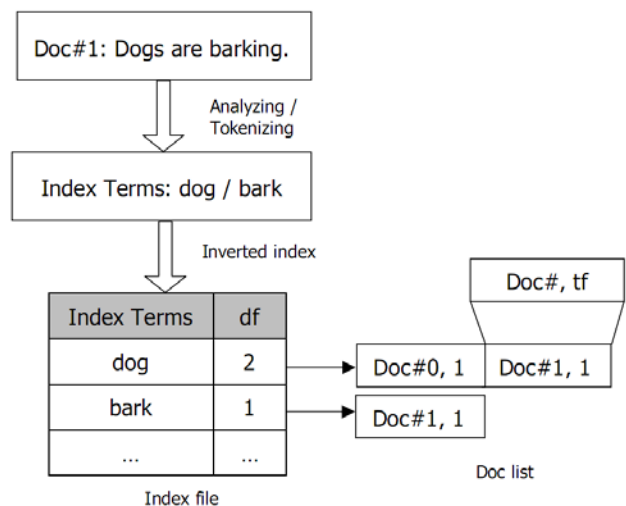
2.1 ทฤษฎี

การค้นหาสารสนเทศ (Information Retrieval: IR) [4] เป็นศาสตร์ของการค้นหาสารสนเทศหรือข้อความที่มีจำนวนมาก ให้ได้มาอย่างรวดเร็ว ตรงกับความต้องการของผู้ใช้มากที่สุด และมีสารสนเทศที่ไม่ตรงความต้องการให้น้อยที่สุด [5] โดยกระบวนการของการค้นหาสารสนเทศเริ่มตั้งแต่ การกำหนดข้อมูลเพื่อระบุขอบเขตข้อความที่ผู้ใช้ต้องการค้นหาจากระบบการประมวลผลข้อความ การสร้างดัชนี การค้นหา การเปรียบเทียบข้อมูล และการนำสารสนเทศที่ให้ผลลัพธ์ตรงตามความต้องการของผู้ค้นอย่างรวดเร็ว ดังภาพที่ 1



ภาพที่ 1: สถาปัตยกรรมของการค้นคืนสารสนเทศ [6]

การสร้างดัชนีแบบแฟ้มผกผัน [5] (Inverted File Index) เป็นการนำเอกสารหรือข้อความมาผ่านการประมวลผลข้อความ (Text Processing) เพื่อสกัดเอาคำที่สำคัญ ๆ มาสร้างดัชนีคำ (Index Term) ซึ่งเมื่อมีการเปลี่ยนแปลงเอกสารหรือการเพิ่มเอกสาร ระบบจะทำการสร้างดัชนีใหม่ เพื่อให้สอดคล้องกับเอกสารใหม่ จากนั้นจึงนำมาเก็บเป็นแฟ้มดัชนีซึ่งมีค่าต่าง ๆ พร้อมทั้งจำนวนเอกสารที่มีค่านั้น ๆ ปรางูอยู่ และแต่ละคำก็จะมีข้อมูลซึ่งระบุรายการหมายเลขของเอกสารพร้อมทั้งจำนวนคำที่ปรากฏอยู่ในเอกสารนั้น ดังภาพที่ 2



Note: df = document frequency, tf = term frequency

ภาพที่ 2: การสร้างดัชนีแบบ Inverted File Index [6]

2.2 งานวิจัยที่เกี่ยวข้อง

ปัญญา [7] ได้นำระบบการค้นคืนข้อมูลพีชสมุนไพรไทย มาช่วยอำนวยความสะดวกแก่ผู้ใช้งานไปใช้ในการรักษาโรค ช่วยทำให้การค้นหาข้อมูลสะดวกรวดเร็วและมีประสิทธิภาพ โดยในการค้นคืนระบบจะรับคำค้นหาจากผู้ใช้งานแล้วนำมาเข้ากระบวนการตัดคำเพื่อนำไปเปรียบเทียบกับข้อมูลในระบบ เมื่อได้ข้อมูลที่ตรงกับคำค้น จะทำการหาคำนำหน้าหนักเพื่อจัดเรียงลำดับของข้อมูล และทำการแสดงผลลัพธ์

วสุวัฒน์ [8] ได้นำเสนอเทคนิคการสร้างดัชนีเพื่อการค้นคืน ด้วยวิธีพาเซิลิสตสแควร์ การใช้วิธีดังกล่าวสำหรับสร้างดัชนี จะทำให้ดัชนีที่ได้มีขนาดเล็กกว่าขนาดดัชนีเดิม โดยในการค้นหาคำจะนำทฤษฎีความคล้ายคลึงเชิงมุมมาหาความสัมพันธ์ระหว่างคำสำคัญที่ต้องการค้นคืน เมื่อลองนำมาทดสอบประสิทธิภาพเทียบกับการสร้างดัชนีที่สร้างจากวิธีการแอลเอสไอ โดยการนำค่าความเที่ยงเฉลี่ย และจับเวลาที่ใช้ในการคำนวณเพื่อสร้างดัชนี ผลที่ได้การสร้างดัชนีตัวแทนเอกสารจากวิธีการพาเซิลิสตสแควร์ ให้ความเที่ยงเฉลี่ยที่มากกว่า จึงแสดงให้เห็นถึงความสามารถในการสร้างดัชนีที่เหมาะสมเมื่อนำมาเทียบกับวิธีการแอลเอสไอซึ่งเป็นวิธีการที่มีการนำเสนออย่างแพร่หลาย

ชูชาติ [9] ได้นำเสนองานวิจัยและพัฒนาโครงสร้างพื้นฐานสารสนเทศอัจฉริยะ ฝ่ายวิจัย และพัฒนาเทคโนโลยีสารสนเทศ (RDI) ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ (NECTEC) ได้ใช้เทคนิคการสร้างดัชนีของลูซีนมาประยุกต์ใช้สำหรับค้นคืนข้อมูล [10] (Lucene for IR Applications) โดยวิธีการสร้างดัชนี จะเริ่มจากการนำเอกสารมาวิเคราะห์และแบ่งเป็นรายการของคำ จากนั้นนำมาสร้างดัชนี (Indexing) และจัดเก็บเป็นไฟล์เอกสาร ซึ่งแต่ละคำจะมีหลายเลขของเอกสาร พร้อมทั้งจำนวนคำที่ปรากฏอยู่ในเอกสารนั้น ๆ ผู้ใช้งานสามารถส่งคิวรีให้กับหน่วยประมวลผลคิวรี หน่วยค้นคืนจะเปรียบเทียบคำที่ค้นคืนกับคำดัชนี แล้วจะจัดลำดับผลลัพธ์การค้นคืนให้กับผู้ใช้งาน

ธีรศักดิ์, สุนทร และวัชรวิไล [11] ได้นำเสนอการประยุกต์ใช้ซอฟต์แวร์ Solr เพื่อการค้นหาข้อมูลภาษาไทยจากหลักสูตรรายวิชา โดยได้นำซอฟต์แวร์ตัวกลางแบบโอเพ่นซอร์ส Apache Solr และโปรแกรมเสริมด้านการวิเคราะห์คำไทย

ได้แก่ วรรณสาร เข้ามาช่วยในการค้นหาและเปรียบเทียบคำค้นภาษาไทย ซึ่งข้อมูลรายวิชาจะจัดเก็บด้วยโครงสร้างข้อมูลแบบ XML หลังจากวิจัยพบว่าการค้นหามีประสิทธิภาพที่ดีในการค้นหาความซ้ำซ้อนของข้อมูลรายวิชา และมีความยืดหยุ่นในการนำไปประยุกต์ใช้พัฒนาเว็บแอปพลิเคชันหรือเว็บเซอร์วิสเหมาะสำหรับการนำไปใช้ในองค์กรที่มีฐานข้อมูลอยู่หลายแหล่งได้เป็นอย่างดี

จากการศึกษางานวิจัยที่เกี่ยวข้องทำให้ผู้วิจัยพบว่า การค้นหาข้อมูลโดยใช้ทฤษฎีการค้นคืน จะช่วยเพิ่มประสิทธิภาพในการค้นหาข้อมูลปริมาณมาก และได้มาซึ่งผลลัพธ์ที่เป็นไปตามความต้องการของผู้ใช้มากขึ้น ภายในระยะเวลาอันสั้น ซึ่งในการเลือกเทคนิคเพื่อนำมาใช้งาน จะขึ้นอยู่กับประเภทหรือรูปแบบของข้อมูล ปริมาณข้อมูลที่เกิดขึ้น ความเร็ว ความถูกต้องของการค้นหา โดยในงานวิจัยครั้งนี้ได้ใช้เทคนิคการตัดคำ และแทนคำด้วยดัชนี โดยเลือกใช้ซอฟต์แวร์ Solr และ Library ของ Microsoft Office Word ซึ่งเป็นซอฟต์แวร์เสรี (free software) และเป็นเครื่องมือที่เหมาะสมกับข้อมูลปัญหา TOT iptv ของบริษัท ทีโอที โดยพิจารณาจากความซับซ้อนของข้อมูล และความต้องการความถูกต้องของผลลัพธ์ที่ไม่สูงมาก อันจะสามารถส่งผลให้การค้นหาข้อมูลเพิ่มประสิทธิภาพมากยิ่งขึ้น

3. วิธีการดำเนินการวิจัย

วิธีการดำเนินงานเพื่อพัฒนาระบบรับแจ้งและค้นคืนปัญหา TOT iptv ทางผู้วิจัยได้กำหนดวิธีการดำเนินงานออกเป็น 4 ขั้นตอน ดังนี้

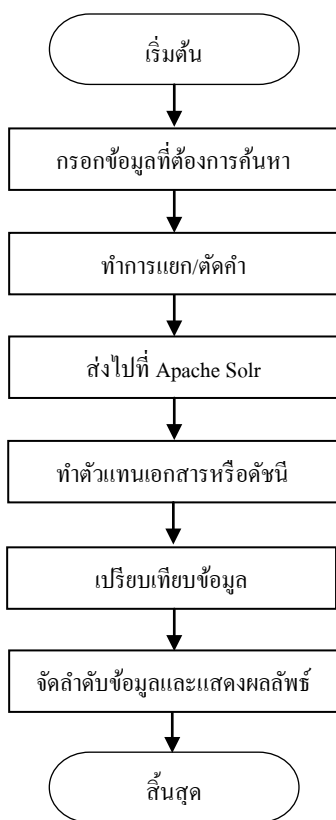
3.1 ศึกษาและวิเคราะห์ปัญหาระบบงาน

ทำการศึกษาและวิเคราะห์ระบบงานเดิม รวมถึงความต้องการของผู้ใช้ เพื่อนำข้อมูลมาค้นหาทฤษฎี หลักการเทคโนโลยี จากเอกสารและผลงานวิจัยที่เกี่ยวข้อง โดยนำข้อมูลที่นำมาเปรียบเทียบข้อดีข้อเสีย และตรวจสอบความเหมาะสมกับสภาพแวดล้อมของระบบเดิม เพื่อนำมาปรับใช้งานต่อไป

3.2 การออกแบบระบบ

เริ่มจากผู้ใช้งานทำการกรอกข้อมูลที่ต้องการค้นคืน และทำการค้นคืน ระบบจะนำข้อมูลมาทำการแยก/ตัดคำ เช่น ปัญหาคุณรายการแล้วดึงกลับหน้าแรก จะถูกตัดคำได้เป็น คุณรายการ,แล้ว

,ตั้ง,กลับ,หน้า,แรก เป็นต้น หลังจาก ตัดคำระบบจะทำการส่งไปที่ Apache Solr เพื่อทำการสร้างตัวแทนเอกสารหรือดัชนี และทำการเปรียบเทียบคั่นคั่นข้อมูลในระบบ ซึ่งกระบวนการตัดคำและสร้างตัวแทนดังกล่าว จะทำในขั้นตอนการนำข้อมูลใหม่เข้าสู่ระบบด้วย หลังจากได้ผลลัพธ์จากการคั่นคั่นจะทำการแสดงผล โดยเรียงลำดับ (ranking) จากการคำนวณค่าน้ำหนักของแต่ละเอกสาร โดยจะคิดจากจำนวนคำที่คั่นเจอเทียบกับจำนวนคำในเนื้อหาเอกสาร เช่น เอกสาร A ถูกคั่นพบ 1 คำ ในจำนวนคำของเอกสารทั้งหมด 20 คำ กับ เอกสาร B ถูกคั่นพบ 2 คำในจำนวนคำของเอกสารจำนวนทั้งหมด 50 เพราะฉะนั้นเอกสาร A จะมีค่าน้ำหนักมากกว่าเอกสาร B เป็นต้น



ภาพที่ 3: แผนผังการทำงานของการค้นคืนในระบบ

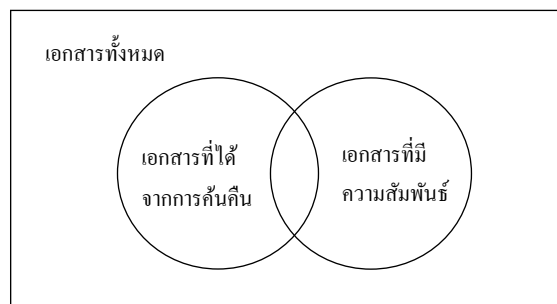
3.3 การพัฒนาระบบ

จากการศึกษาและวิเคราะห์ระบบงาน ได้ทำการพัฒนาระบบตามที่ได้ออกแบบไว้ โดยผู้วิจัยได้เลือกใช้โปรแกรมภาษา C# (ASP.NET) ที่อยู่ในชุดของโปรแกรม Microsoft Visual Studio 2010 Express และได้ใช้ชุดโปรแกรมดังกล่าว

ร่วมกับ JavaScript ในการสร้างเว็บเพจ และหน้าจอส่วนที่ติดต่อกับผู้ใช้งาน เพื่อให้ระบบสามารถทำงาน และตอบสนองการใช้งานของผู้ใช้ได้อย่างมีประสิทธิภาพมากขึ้น ในส่วนฐานข้อมูลผู้วิจัยได้เลือกใช้โปรแกรม Microsoft SQL Server 2008 Express Edition ซึ่งจะติดต่อผ่านทาง ADO.NET ส่วนการตัดคำเลือกใช้ Library ของ Microsoft Office Word และการทำตัวแทนเอกสารหรือดัชนีเลือกใช้ Apache Solr

3.4 การทดสอบระบบ

สำหรับการทดสอบประสิทธิภาพของระบบนั้นจะใช้การวัดค่าความแม่นยำ (Precision) และค่าการเรียกคืน (Recall) [12] เพื่อใช้ตรวจสอบว่าระบบที่พัฒนาขึ้นสามารถค้นคืนได้อย่างมีประสิทธิภาพ [6] ดังภาพที่ 4



ภาพที่ 4: แสดงเซตของเอกสารที่ได้จากการค้นคืน

ความแม่นยำ คือ อัตราส่วนระหว่างเซตของเอกสารที่ได้จากการค้นคืนดังสมการ

$$\text{Precision} = \frac{\text{Number of relevant documents retrieved}}{\text{Total number of documents retrieved}} \quad (1)$$

จากสมการที่ 1 คำนวณหาค่าความแม่นยำจากเอกสารที่เกี่ยวข้องและถูกค้นคืนทั้งหมดหารกับจำนวนเอกสารทั้งหมดที่ถูกค้นคืน

ค่าการเรียกคืน คือ อัตราส่วนของเอกสารที่เกี่ยวข้องที่ถูกค้นคืนจากจำนวนเอกสารที่เกี่ยวข้องทั้งหมดดังสมการ

$$\text{Recall} = \frac{\text{Number of relevant documents retrieved}}{\text{Total number of relevant documents}} \quad (2)$$

จากสมการที่ 2 คำนวณหาค่าการเรียกคืนได้จากเอกสารที่ค้นคืนและเกี่ยวข้องกับทั้งหมดหารกับจำนวนเอกสารที่เกี่ยวข้องทั้งหมด

4. ผลการดำเนินงาน

4.1 ผลการพัฒนาระบบ

ผลการดำเนินงานของงานวิจัยในครั้งนี้สามารถพัฒนาระบบการรับแจ้งและค้นคืนปัญหา TOT iptv ได้ผลลัพธ์ดังตัวอย่างต่อไปนี้



ภาพที่ 5: หน้าจอแสดงข้อมูลการค้นหของ Apache Solr



ภาพที่ 6: หน้าจอค้นหาข้อมูลปัญหา

4.2 ผลการทดสอบประสิทธิภาพ

ในการทดสอบระบบเพื่อวัดประสิทธิภาพการค้นคืน ได้ทำการทดสอบการค้นหาค่าโดยวัดประสิทธิภาพด้วยค่าความแม่นยำ (Precision) และค่าการเรียกคืน (Recall) ซึ่งทดสอบจากคำค้นหาจำนวน 50 คำ โดยตัวอย่างของผลการทดสอบการค้นหาคำสำคัญ มีรายละเอียดดังนี้

ตารางที่ 1: ตัวอย่างผลการทดสอบการค้นหาคำสำคัญ เพื่อวัดค่าประสิทธิภาพของระบบ

คำที่ค้นหา	เอกสาร			Precision	Recall
	คำค้น	เกี่ยวข้อง	ค้นคืนและเกี่ยวข้อง		
กล่อง STB	33	37	27	0.82	0.73
อินเทอร์เน็ต	28	23	18	0.64	0.78
หน้าจอดี	31	34	22	0.71	0.65
ภาพกระตุก	23	18	15	0.65	0.83
iPad	29	35	21	0.72	0.60
รหัส -7	46	40	35	0.76	0.87
ภาพไม่ชัด	31	27	21	0.67	0.77
มีภาพไม่มีเสียง	25	27	19	0.76	0.70
รหัส -2	17	21	16	0.94	0.76
จอหมุน	33	31	26	0.78	0.87

จากตารางที่ 1 สรุปผลการวัดประสิทธิภาพการค้นคืน จากการทดสอบโดยใช้คำค้นหาจำนวน 50 คำ พบว่าระบบมีประสิทธิภาพการค้นคืนแม่นยำเฉลี่ย ร้อยละ 74 และค่าการเรียกคืนเฉลี่ย ร้อยละ 76 ดังนั้นระบบที่พัฒนาขึ้นมีประสิทธิภาพในการค้นคืนอยู่ในระดับดี

5. สรุปผลและข้อเสนอแนะ

ระบบปรับแก้และค้นคืนปัญหา TOT iptv ได้ใช้เทคนิคการตัดคำและสร้างดัชนีในแต่ละเอกสาร โดยประยุกต์โซลาร์ [13] (Solr) ซึ่งมีพื้นฐานมาจากระบบสืบค้นลูซีน [14] เพื่อช่วยในการเปรียบเทียบคำที่มีความสอดคล้องกันในแต่ละเอกสารในปริมาณจำนวนมาก ให้ได้รับผลลัพธ์ของการค้นหาที่เป็นไปตามความต้องการของผู้ใช้มากขึ้น ภายในระยะเวลาที่ใช้น้อยลง

จากผลการทดสอบระบบได้ค่าความแม่นยำ (Precision) ร้อยละ 74 และค่าการเรียกคืน (Recall) ร้อยละ 76 ทำให้ระบบมีประสิทธิภาพในการค้นคืนเป็นที่น่าพอใจ อย่างไรก็ตามในอนาคตหากสามารถพัฒนาการค้นคืนให้รองรับคำที่มีความหมายคล้ายคลึงกันหรือแม้กระทั่งคำพ้องเสียงให้สามารถค้นคืนได้ จะทำให้การค้นคืนคำที่ต้องการมีความยืดหยุ่นและได้ผลลัพธ์ที่มีประสิทธิภาพมากขึ้น นอกจากนี้การพัฒนาระบบในรูปแบบแอปพลิเคชันที่สามารถใช้งานไบนารีโทรศัพท์เคลื่อนที่แบบสมาร์ทโฟนหรืออุปกรณ์แท็บเล็ตคอมพิวเตอร์ ก็จะช่วยอำนวยความสะดวกในการใช้งาน และตอบสนองการใช้งานแก่ผู้ใช้ได้หลากหลายรูปแบบ ทำให้ระบบมีประสิทธิภาพในการค้นคืนเป็นที่น่าพอใจ

เอกสารอ้างอิง

- [1] การประยุกต์ในเทคโนโลยีสารสนเทศกับธุรกิจทิวทัศน์. “การประยุกต์ใช้เทคโนโลยีกับบริษัท True Visions”, สืบค้นวันที่ 8 มกราคม 2557 จาก <http://bewtayed.blogspot.com/2011/01/blog-post.html>
- [2] วิงค์รอง ศรีสวัสดิ์, “ระบบช่วยเหลือในการค้นคืนสารสนเทศ”, *วารสารวิทยบริการ* ปีที่ 15 ฉบับที่ 2-3 พ.ศ.-ช.ค. 2547 หน้า 24.
- [3] อรรถกร จันทร์ทอง, “Twitter หันมาใช้ Apache Lucene เป็นระบบการค้นหา”, สืบค้นวันที่ 12 มกราคม 2557 จาก <http://www.blognone.com/news/19059/>
- [4] ศิริรัตน์ ศิรินานนท์, “การค้นคืนสารสนเทศโดยใช้กฎความสัมพันธ์ร่วมกับผลสะท้อนกลับจากผู้ใช้”, *คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย*, 2549.
- [5] R. Baeza-Yates and B. Ribeiro-Neto, “Moder Information Retrieval”, ACM Press, Addison Wesley, 1999.
- [6] อัครพล วุฒิเกรียงไกรผล, “การเพิ่มประสิทธิภาพระบบค้นคืนสารสนเทศรูปแบบเอกสารอิเล็กทรอนิกส์ โดยใช้พีชเชอจิก”, *คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*, 2554.
- [7] ปัญญา คนทนต์, “ระบบค้นคืนข้อมูลพีชสมุนไพรไทย”, *คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*, 2554.
- [8] วสุวัฒน์ ทิพย์สังวาลย์, “เทคนิคการสร้างดัชนีเพื่อการค้นคืนด้วยวิธีพาเซิลีสสแควร์”, *คณะวิทยาศาสตร์ มหาวิทยาลัยเชียงใหม่*, 2553.
- [9] ชูชาติ หุไชยะศักดิ์, “โปรแกรมสำหรับพัฒนาระบบค้นคืนสารสนเทศภาษาไทย”, National and Computer Technology Center (NECTEC). สืบค้นวันที่ 15 มกราคม 2557 จาก <http://sansarn.com>
- [10] ชีรศักดิ์ เพ็ญมูขอ สุนทร วิทูสรพจน์ และวัชรวิดี ตั้งคุปตานนท์, “การประยุกต์ใช้ซอฟต์แวร์ Solr เพื่อการค้นหาข้อมูลภาษาไทยจากหลักสูตรรายวิชา”, สืบค้นวันที่ 22 ธันวาคม 2556 จาก http://www.lib.pt.tsu.ac.th/rpms/files_upload/news/%5B2012-12-12-14-53-16%5Dการประยุกต์ใช้ซอฟต์แวร์%20Solr.pdf
- [11] E. Hatcher and O. Gospo, “Lucene in Action”, Manning Publications, 2005.
- [12] ธนพล ดันดิศรีปรีชา, “การใช้ออนโทโลยีเพื่อการค้นคืนคำพิพากษาศาลฎีกา”, *คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์*, 2553.
- [13] The Apache Software Foundation. “Apache Solr”, สืบค้นวันที่ 12 มกราคม 2557 จาก <http://lucene.apache.org/solr/>
- [14] The Apache Software Foundation. “Apache Lucene”, สืบค้นวันที่ 12 มกราคม 2557 จาก <http://lucene.apache.org/>

การสร้างแบบจำลองการพยากรณ์ราคากองทุนรวม โดยใช้เหมืองข้อมูลและโครงข่ายประสาทเทียม

NAV Forecasting Model Using Data Mining and Artificial Neural Network

ธารารัตน์ อุดมชัยบรรเจิด (Thararat Udomchaibanjerd)¹ และผุสดี บุญรอด (Pudsadee Boonrawd)²

^{1,2}ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
thararat.udom@gmail.com¹, pudsadee@kmutnb.ac.th²

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อสร้างแบบจำลองสำหรับพยากรณ์ราคากองทุนรวมตราสารหนี้ โดยใช้หลักการของโครงข่ายประสาทเทียมแบบหลายชั้น ลักษณะการเรียนรู้แบบแพร่กระจายย้อนกลับ เพื่อให้เห็นแนวโน้มของราคากองทุนรวม และช่วยสนับสนุนการตัดสินใจลงทุนในกองทุนรวมได้อย่างเหมาะสม โดยใช้ข้อมูลผลิตภัณฑ์มวลรวมในประเทศ ข้อมูลอัตราดอกเบี้ยนโยบาย ข้อมูลดัชนีราคาหุ้นตลาดหลักทรัพย์แห่งประเทศไทย ข้อมูลดัชนีราคา SET100 และข้อมูลดัชนีตราสารหนี้ ตั้งแต่ปี พ.ศ. 2551 ถึง พ.ศ. 2556 แบบจำลองนี้มีข้อมูลนำเข้า 8 ตัวแปร และมีผลลัพธ์ 1 ตัวแปร จากนั้นทดลองหาจำนวนนิวรอนในชั้นซ่อนในรูปแบบที่แตกต่างกัน เพื่อหาแบบจำลองที่ให้ค่าความคลาดเคลื่อนเฉลี่ยยกกำลังสองน้อยที่สุด ซึ่งผลการทดลองพบว่าโครงข่ายประสาทเทียมที่เหมาะสมที่สุดมีจำนวนนิวรอนในชั้นนำเข้า จำนวนนิวรอนในชั้นซ่อน และจำนวนนิวรอนในชั้นแสดงผลเท่ากับ 8-10-1 ตามลำดับ ผลการทดสอบพบว่าการพยากรณ์มีค่าความคลาดเคลื่อนเฉลี่ยยกกำลังสองน้อยที่สุดเท่ากับ 0.0062 สรุปได้ว่าแบบจำลองมีความเหมาะสมสำหรับพยากรณ์ราคากองทุนรวมตราสารหนี้ และสามารถนำไปประยุกต์ใช้งานได้จริง

คำสำคัญ: โครงข่ายประสาทเทียม การพยากรณ์ ราคากองทุน

Abstract

This research aims to develop a NAV forecasting model using Artificial Neural Networks (ANN) with multi-layer perceptions from the Back Propagations algorithm to see the trend of mutual funds and support

the decision to invest in mutual funds appropriately. In this paper, GDP, Policy Interest Rate, SET Index, SET100 Index and Bond Index data during 2008 to 2013 are used with 8 inputs and 1 output by a neural network then examined to find hidden layers to organize the minimum Mean Square Error of the neural network model. According to the results, ANN structure 8-10-1 was selected to be optimal as provided the best Mean Square Error of 0.0062. In conclusion, the NAV forecasting model clearly achieved all objectives successfully.

Keywords: Artificial neural Network, Forecast, NAV

1. บทนำ

จากสถานการณ์ในปัจจุบันของสภาวะดอกเบี้ยเงินฝากของสถาบันการเงินมีแนวโน้มลดลงอย่างต่อเนื่อง ซึ่งส่งผลกระทบต่อผู้ที่มีเงินฝากที่ต้องการผลตอบแทนให้เพียงพอกับความ ต้องการ ซึ่งอาจจะต้องนำเงินออมไปลงทุนในรูปแบบอื่น เพื่อให้ได้ผลตอบแทนมากขึ้นกว่าการออมเงินในรูปแบบเงินฝากเพียงอย่างเดียว ซึ่งทางเลือกในการลงทุนมีหลากหลายตามผลตอบแทนที่ต้องการ กองทุนรวมเป็นทางเลือกหนึ่งในการลงทุนที่ได้รับความนิยมเป็นอย่างมากในกลุ่มผู้ลงทุนรายย่อย [1] ที่มีความประสงค์จะนำเงินมาลงทุนในตลาดเงินและตลาดทุน แต่อาจคิดขัดหลาย ๆ ประการที่ทำให้การลงทุนด้วยตนเองไม่สามารถได้ผลตอบแทนตามเป้าหมายที่ต้องการ กองทุนรวมจึงเป็นเครื่องมือในการลงทุนที่มีประสิทธิภาพ เพราะมีการจัดการลงทุนอย่างเป็นระบบ เพื่อให้การลงทุนได้รับผลตอบแทนที่ดีที่สุดภายใต้ระดับความเสี่ยงที่ผู้ลงทุนยอมรับได้

เนื่องจากกองทุนรวมมีหลากหลายประเภท และวัตถุประสงค์ในการลงทุนของแต่ละกองทุนแตกต่างกัน ความเสี่ยงและผลตอบแทนที่จะได้รับของแต่ละกองทุนจึงแตกต่างกันด้วย ทำให้ผู้ลงทุนตัดสินใจยากเพราะผู้ลงทุนส่วนใหญ่ยังไม่มีการพิจารณาถึงปัจจัยดังกล่าวที่ดีพอ จึงอาจจะทำให้ได้รับผลตอบแทนไม่ดีเท่าที่ควรหรืออาจจะขาดทุนจากการเลือกลงทุนในกองทุนที่ไม่เหมาะสมกับตนเอง จากปัญหาดังกล่าวในส่วนของงานวิจัยนี้จึงได้ออกแบบและพัฒนาแบบจำลองการพยากรณ์ราคากองทุนรวมตราสารหนี้ โดยใช้เหมือนข้อมูลและโครงข่ายประสาทเทียมแบบเปอร์เซพตรอนหลายชั้น (MLP: Multi Layer Perceptron) และใช้ลักษณะการเรียนรู้แบบแพร่กระจายย้อนกลับ (Back Propagation Learning) ที่เหมาะสมสำหรับพยากรณ์ราคากองทุนรวม เพื่อช่วยสนับสนุนการตัดสินใจลงทุนในกองทุนรวมที่มีนโยบายการลงทุนที่เหมาะสมกับความต้องการของนักลงทุนในระดับอัตราผลตอบแทนและความเสี่ยงที่ยอมรับได้ ซึ่งเนื้อหาของงานวิจัยนี้แบ่งออกเป็น 5 หัวข้อหลัก คือ บทนำ ทฤษฎีและงานวิจัยที่เกี่ยวข้อง วิธีการดำเนินการวิจัย ผลการดำเนินการวิจัย และสรุปผลการวิจัยและข้อเสนอแนะ

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ความรู้เกี่ยวกับกองทุนรวม

กองทุนรวมเป็นกองทรัสต์สินทรัพย์ที่จัดตั้งและบริหารโดยบริษัทหลักทรัพย์จัดการกองทุนหรือเรียกว่า บลจ. มีวัตถุประสงค์ที่นำเงินที่ได้จากการจำหน่ายหน่วยลงทุนแก่ผู้ลงทุนไปลงทุนในหลักทรัพย์ประเภทต่าง ๆ ที่ให้ผลตอบแทนและมีความเสี่ยงในการลงทุนแตกต่างกัน เพื่อให้ผู้ลงทุนนั้นได้รับผลตอบแทนในรูปแบบเงินปันผลหรือกำไรจากการขายหน่วยลงทุน [1] โดยกองทุนแบ่งออกเป็น 2 ประเภท คือ กองทุนปิดเป็นกองทุนที่มีหน่วยลงทุนคงที่และมีกำหนดอายุโครงการแน่นอน และกองทุนเปิดเป็นกองทุนรวมที่สามารถเพิ่มหรือลดหน่วยลงทุนได้ ซึ่งมีหลายประเภทตามระดับความเสี่ยงในการลงทุน หนึ่งในนั้นคือกองทุนรวมตราสารแห่งหนี้ (General Fixed Income Fund) เป็นกองทุนรวมที่มีนโยบายการลงทุนในหรือมีไว้เฉพาะเงินฝากหรือหลักทรัพย์หรือทรัพย์สินอื่น หรือการหาดอกผลโดยวิธีอื่น

ตามที่สำนักงาน ก.ล.ต. กำหนดหรือให้ความเห็นชอบให้กองทุนประเภทดังกล่าวลงทุนได้

โดยทั่วไปแล้วกองทุนรวมตราสารแห่งหนี้ มีความเสี่ยงน้อยกว่ากองทุนรวมที่มีนโยบายลงทุนในตราสารทุน จึงเหมาะสำหรับผู้ลงทุนที่ยอมรับความเสี่ยงได้น้อยกว่า [1] และในแต่ละวันกองทุนรวมจะทำการคำนวณมูลค่าทรัพย์สินของกองทุนตามราคาตลาด (Mark to Market) เพื่อให้สะท้อนถึงมูลค่าที่เป็นจริงตามสถานะตลาดที่ได้เปลี่ยนแปลงไป โดยมูลค่าทรัพย์สินสุทธิ (NAV: Net Asset Value) คือ มูลค่าทรัพย์สินทั้งหมดของกองทุนรวมตลอดจนผลประโยชน์ต่าง ๆ ที่กองทุนรวมได้รับจากการลงทุน ณ เวลาขณะใดขณะหนึ่งหักออกด้วยค่าใช้จ่ายและหนี้สินของกองทุนรวมนั้น

2.2 เหมือนข้อมูล (Data Mining)

เหมือนข้อมูล คือ กระบวนการในการค้นหาความสัมพันธ์รูปแบบและแนวโน้มของข้อมูล หรือเป็นการค้นหาวิเคราะห์หรือสร้างองค์ความรู้ใหม่จากข้อมูลที่ถูกเก็บรวบรวมไว้เป็นจำนวนมาก โดยจะใช้วิธีการทางสถิติหรือวิธีการทางคณิตศาสตร์ ซึ่งมี 5 รูปแบบ [2] คือ

2.2.1 Association Rule การค้นหาหาความสัมพันธ์หรือความเชื่อมโยงของข้อมูลสองชุดหรือมากกว่าขึ้นไปไว้ด้วยกัน

2.2.2 Classification and Prediction การจำแนกประเภทและการทำนาย ซึ่งในงานวิจัยนี้ได้ทำการค้นหาแบบจำลองโครงข่ายประสาทเทียมที่เหมาะสมในการพยากรณ์ราคากองทุนรวม

2.2.3 Cluster Analysis การจัดกลุ่มข้อมูลที่มีความคล้ายคลึงกันมากที่สุด

2.2.4 Outlier Analysis การหาค่าผิดปกติที่เกิดขึ้น

2.2.5 Trend and evolution Analysis การวิเคราะห์แนวโน้ม

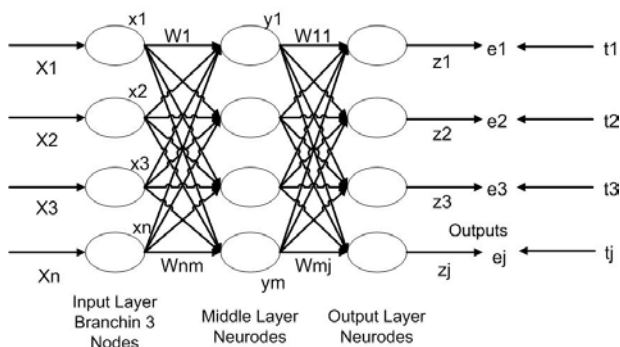
การทำเหมือนข้อมูลถูกนำไปประยุกต์ใช้ในงานหลายประเภท ทั้งในด้านธุรกิจที่ช่วยในการตัดสินใจของผู้บริหารด้านวิทยาศาสตร์และการแพทย์ รวมถึงด้านเศรษฐกิจและสังคม

2.3 โครงข่ายประสาทเทียม

โครงข่ายประสาทเทียม (Artificial Neural Network) คือแบบจำลองทางคณิตศาสตร์ สำหรับประมวลผลสารสนเทศเพื่อจำลองการทำงานของโครงข่ายประสาทในสมองมนุษย์ ให้มีความสามารถในการเรียนรู้การจดจำรูปแบบ

(Pattern Recognition) และการอุปมาความรู้อื่น (Knowledge Deduction) เช่นเดียวกับความสามารถที่มีในสมองมนุษย์ โดยสามารถนำมาใช้ในการพยากรณ์ (Forecasting) [3] ซึ่งเป็นการคาดคะเนลักษณะต่าง ๆ หรือเป็นการประเมินสิ่งที่จะเกิดขึ้นในอนาคตด้วยการคาดการณ์ล่วงหน้า โดยใช้ข้อมูลที่ผ่านมานในอดีต การพยากรณ์จึงมีบทบาทสำคัญในทุก ๆ ภาคส่วนที่จะช่วยให้ได้ข้อมูลในอนาคตเพื่อประกอบการวางแผนสนับสนุนการตัดสินใจในเรื่องต่าง ๆ

โครงข่ายประสาทเทียมแบบเพอร์เซปตรอนหลายชั้น หรือโครงข่ายประสาทเทียมแบบแพร่กระจายย้อนกลับ ซึ่งเป็นโครงข่ายประสาทเทียมแบบต้องมีผู้สอน (Supervised Learning) โดยสถาปัตยกรรมของโครงข่ายประสาทเทียมแบบแพร่กระจายย้อนกลับมีลักษณะหลัก ๆ ดังนี้ คือ โครงข่ายประสาทเทียมจะประกอบด้วยชั้นต่าง ๆ ได้แก่ ชั้นนำเข้า (Input Layer) ชั้นซ่อน (Hidden Layer) และชั้นส่งออก (Output Layer) และการเชื่อมต่อระหว่างชั้นต่าง ๆ นั้น ทุก ๆ โหนดในชั้นนำเข้าจะส่งสัญญาณไปยังทุก ๆ โหนดในชั้นซ่อนชั้นแรกและทุก ๆ โหนดในชั้นซ่อนชั้นแรกจะส่งสัญญาณไปยังทุก ๆ โหนดในชั้นซ่อนชั้นสุดท้ายจะส่งสัญญาณไปยังทุก ๆ โหนดในชั้นส่งออกดังภาพที่ 1



ภาพที่ 1: โครงข่ายประสาทเทียมแบบแพร่กระจายย้อนกลับ

ขั้นตอนวิธีการแพร่กระจายย้อนกลับ โดยฟังก์ชันที่นิยมใช้ในโครงข่ายประสาทเทียมแบบเพอร์เซปตรอนหลายชั้นสอนได้แก่ Sigmoid Function โดยมี 2 ขั้นตอนดังนี้ [3] การคำนวณไปข้างหน้า (Forward Computation) คือ การนำข้อมูลนำเข้า (Input Data) ส่งผ่านไปข้างหน้าตามลำดับชั้น ซึ่งมีการรวมข้อมูล Input กับค่า Weight เข้าด้วยกันและส่งผ่านฟังก์ชันแปลงค่า (Transfer Function) เพื่อทำการแปลงให้เป็นผลลัพธ์ภายในชั้นแฝง จากนั้นจึงนำผลลัพธ์ที่ได้จากการคำนวณ

(Output) มาเปรียบเทียบกับข้อมูลจริง (Target Output) และทำการคำนวณค่าความคลาดเคลื่อนทั้งหมดที่เกิดขึ้น (Total Error) เมื่อผลลัพธ์ที่คำนวณได้มีความคลาดเคลื่อนจะเริ่มขั้นตอนในการคำนวณย้อนกลับ (Backward Computation) ซึ่งเริ่มจากการปรับค่า Weight ระหว่างชั้นผลลัพธ์กับชั้นแฝงชั้นสุดท้ายและย้อนกลับไปจนถึงชั้นรับข้อมูล

2.4 งานวิจัยที่เกี่ยวข้อง

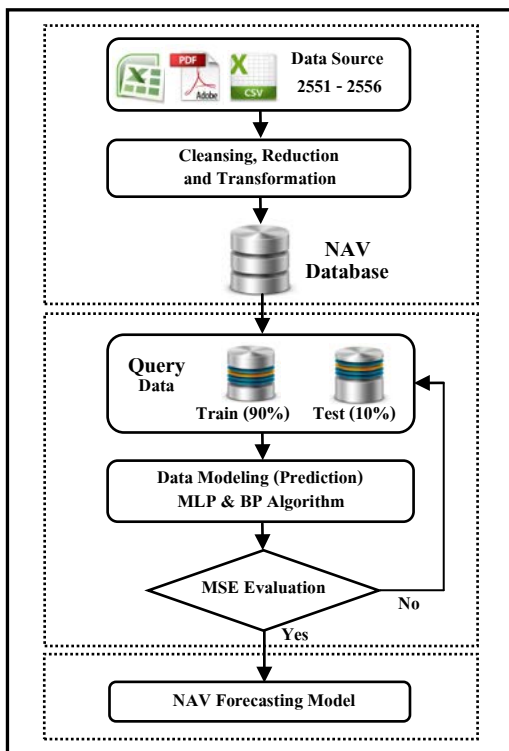
งานวิจัยที่เกี่ยวกับการพยากรณ์ส่วนใหญ่ก็จะนำเหมืองข้อมูลเข้ามาช่วยในการวิเคราะห์ข้อมูลเพื่อหาองค์ความรู้ใหม่ หรือความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลให้เกิดประโยชน์สูงสุด และในส่วนของอัลกอริทึมที่นิยมนำมาใช้ในการพยากรณ์ที่ให้ผลลัพธ์ที่ผิดพลาดน้อยที่สุด ก็คือ โครงข่ายประสาทเทียม เพราะสามารถวิเคราะห์ข้อมูลได้ทั้งลักษณะที่เป็นเชิงเส้นและไม่เป็นเชิงเส้นได้ดี โดยมีงานวิจัยที่เกี่ยวกับการลงทุนทั้งในหุ้นและกองทุน [4] ศึกษารูปแบบในอนุกรมเวลาของราคาหุ้นด้วย Closing Price Time Series เพื่อทำนายแนวโน้มในอนาคตของตลาดหุ้นโดยใช้ราคาหุ้นของ Iran Tractor Manufacturing Company [5] พัฒนาแบบจำลองการทำนายหุ้น โดยทดลองกับข้อมูลจาก Shanghai Stock [6] นำเสนอ Modified Neural Network Forecasting Model และพัฒนา Intelligent Mining System โดยระบบสามารถคาดการณ์แนวโน้มในอนาคตของสัญญาณซื้อและขายของ Shanghai Composite Index และในหลายงานวิจัยก็ได้้นำเอาวิธีการต่าง ๆ มาเปรียบเทียบกับกัน [7] นำตัวแบบ ARIMA, ANN และตัวแบบผสม ARIMA-ANN มาทำการเปรียบเทียบความถูกต้องในการพยากรณ์อนุกรมเวลาราคาปิดตลาดของหุ้น PTT [8] คาดการณ์การเคลื่อนไหวของดัชนีตลาดหลักทรัพย์แห่งประเทศไทย (SET Index) [9] สํารวจข้อดีและข้อเสียของ Smooth Support Vector Regression (SSVR) และ Back Propagation Network (BPN) ในการคาดการณ์ผลการดำเนินงานกองทุนรวมของ Taiwan's Equity Fund

3. วิธีการดำเนินการวิจัย

การดำเนินการวิจัยในการออกแบบและพัฒนาแบบจำลองการพยากรณ์ราคากองทุนรวมโดยใช้เหมืองข้อมูลและโครงข่ายประสาทเทียมจะมีขั้นตอนต่าง ๆ ดังนี้

3.1 การรวบรวมและวิเคราะห์ข้อมูล

การเก็บรวบรวมข้อมูลและวิเคราะห์ปัจจัยที่มีผลกระทบต่อราคากองทุนรวม ได้แก่ ข้อมูลผลิตภัณฑ์มวลรวมในประเทศ ข้อมูลอัตราดอกเบี้ยนโยบาย ข้อมูลดัชนีราคาหุ้นตลาดหลักทรัพย์แห่งประเทศไทย ข้อมูลดัชนีราคา SET100 และข้อมูลดัชนีตราสารหนี้ ซึ่งเป็นข้อมูลที่มีการเก็บรวบรวมเป็นรายวัน ตั้งแต่ กรกฎาคม พ.ศ. 2551 ถึง มิถุนายน พ.ศ. 2556 ทั้งหมด 5 ปี เพื่อนำมาพยากรณ์ราคากองทุนรวมตราสารหนี้ของ 3 บริษัทหลักทรัพย์จัดการกองทุนในประเทศไทยที่มีส่วนแบ่งการตลาดของกองทุนรวมใหญ่เป็น 3 ลำดับแรก (ข้อมูล ณ วันที่ 30 สิงหาคม 2556) ได้แก่ บริษัทหลักทรัพย์จัดการกองทุนกสิกรไทย จำกัด บริษัทหลักทรัพย์จัดการกองทุนไทยพาณิชย์ จำกัด และบริษัทหลักทรัพย์จัดการกองทุนกรุงไทย จำกัด (มหาชน) ตามกรอบแนวคิดการพยากรณ์ราคากองทุนรวมดังภาพที่ 2



ภาพที่ 2: กรอบแนวคิดการพยากรณ์ราคากองทุนรวม

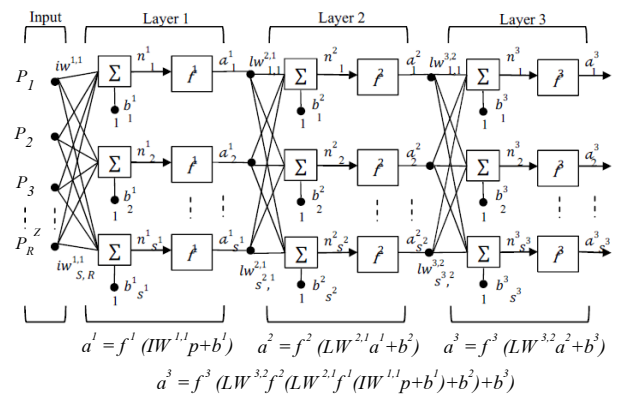
3.2 การเตรียมข้อมูลสำหรับสร้างแบบจำลอง

การนำเอาข้อมูลหลากหลายรูปแบบที่เก็บรวบรวมมาทำการคัดกรองข้อมูล (Data Cleaning) โดยกำจัดข้อมูลที่ไม่จำเป็นออกไปหรือใส่ค่าที่หายไปของข้อมูลให้ครบถ้วน จากนั้นทำการลดรูป (Data Reduction) และแปลงข้อมูล

(Data Transformation) ให้อยู่ในรูปแบบที่เหมาะสมสำหรับการนำไปใช้สร้างแบบจำลองการพยากรณ์ รวมถึงการแบ่งชุดข้อมูลออกเป็น 2 ชุด คือ ชุดข้อมูลสำหรับฝึกสอน (Training) 90% และชุดข้อมูลสำหรับทดสอบ (Testing) 10%

3.3 การสร้างแบบจำลอง

งานวิจัยนี้ได้พัฒนาแบบจำลองการพยากรณ์ราคากองทุนรวมเป็นรายวัน โดยมีข้อมูลนำเข้า 8 ตัวแปร ได้แก่ ปี เดือน รหัสกองทุน ข้อมูลผลิตภัณฑ์มวลรวมในประเทศ ข้อมูลอัตราดอกเบี้ยนโยบาย ข้อมูลดัชนีราคาหุ้นตลาดหลักทรัพย์แห่งประเทศไทย ข้อมูลดัชนีราคา SET100 และข้อมูลดัชนีตราสารหนี้ และมีผลลัพธ์ 1 ตัวแปร คือ ราคากองทุน จากนั้นทดลองหาจำนวนนิรอนในชั้นซ่อนในรูปแบบที่แตกต่างกันจำนวน 10 รูปแบบ เพื่อออกแบบโครงข่ายประสาทเทียมที่ทำให้ค่าความคลาดเคลื่อนเฉลี่ยยกกำลังสองน้อยที่สุด โดยกำหนดจำนวนรอบในการฝึกสอนเท่ากับ 500 รอบ อัตราการเรียนรู้เท่ากับ 0.1 และค่าโมเมนตัม เท่ากับ 0.2 ซึ่งการคำนวณดังภาพที่ 3 [10]



ภาพที่ 3: โครงสร้างโครงข่ายประสาทเทียมแบบหลายชั้น

โดยที่ $W^{1,1}$ คือ ค่าถ่วงน้ำหนักที่เชื่อมระหว่างชั้นข้อมูลนำเข้าและชั้นซ่อนที่ 1

$W^{2,1}$ คือ ค่าถ่วงน้ำหนักที่เชื่อมระหว่างชั้นซ่อนที่ 1 และชั้นซ่อนที่ 2

$W^{3,2}$ คือ ค่าถ่วงน้ำหนักที่เชื่อมระหว่างชั้นซ่อนที่ 2 และชั้นผลลัพธ์

b^1 คือ ค่าไบแอสในชั้นซ่อนที่ 1

b^2 คือ ค่าไบแอสในชั้นซ่อนที่ 2

b^3 คือ ค่าไบแอสในชั้นผลลัพธ์

ซึ่งค่าถ่วงน้ำหนัก (W) และไบแอส (b) สามารถปรับค่าได้ตามการเรียนรู้ และผลลัพธ์จะขึ้นอยู่กับฟังก์ชันการถ่ายโอนดังสมการที่ 1

$$N = W^{11}X^1 + W^{12}X^2 + \dots + W^{1R}X^R + b \quad (1)$$

เขียนให้อยู่ในรูปเมตริกซ์ได้ดังสมการที่ 2

$$n = W^X + b \quad (2)$$

โดยที่เมตริกซ์ W เป็น Row Matrix มีเพียงแถวเดียว จะได้ผลลัพธ์ (Y) ดังสมการที่ 3

$$Y = f(W^X + b) \quad (3)$$

ถ้าผลลัพธ์ที่ได้ไม่ตรงกับค่าเป้าหมาย จะกลับไปทำการเรียนรู้ใหม่โดยปรับค่าถ่วงน้ำหนักและไบแอสใหม่ส่งต่อไปยังฟังก์ชันถ่ายโอนจนกว่าผลลัพธ์ที่ได้จะเข้าใกล้ค่าเป้าหมายมากที่สุด

3.4 การวัดค่าความแม่นยำของแบบจำลอง

ทำการวัดจากแบบจำลองที่ให้ค่าความคลาดเคลื่อนเฉลี่ยยกกำลังสองน้อยที่สุด (MSE: Mean Square Error) ดังสมการที่ 4

$$MSE = \frac{1}{N} \sum_{l=1}^N (e_l)^2 = \frac{1}{N} \sum_{l=1}^N (t_l - a_l)^2 \quad (4)$$

โดยที่ t_l คือ ค่าจริง

a_l คือ ค่าพยากรณ์

e_l คือ ค่าความคลาดเคลื่อน

N คือ จำนวนค่าพยากรณ์

3.5 การวิเคราะห์และออกแบบระบบงาน

นำแบบจำลองที่เหมาะสมที่สุดมาออกแบบและพัฒนาเป็นระบบพยากรณ์ราคากองทุนรวมในรูปแบบของเว็บแอปพลิเคชัน ซึ่งในส่วนของ การวิเคราะห์และออกแบบระบบงานได้แบ่งผู้ใช้งานออกเป็น 2 กลุ่ม ได้แก่ กลุ่มที่ 1 ผู้ใช้งานทั่วไปสามารถเรียกดูข้อมูลกองทุนรวมและราคาของกองทุนรวมย้อนหลังได้ และสามารถทำแบบทดสอบประเมินความเสี่ยงในการลงทุน เพื่อสืบค้นกองทุนรวมที่มีความเหมาะสมกับผู้ใช้งานในระดับความเสี่ยงที่ยอมรับได้ รวมถึงสามารถพยากรณ์ราคากองทุนรวมได้ล่วงหน้า และเปรียบเทียบผลตอบแทนของแต่ละกองทุนรวมได้ กลุ่มที่ 2 ผู้ดูแลระบบสามารถจัดการกับข้อมูลส่วนต่าง ๆ ของระบบได้

4. ผลการดำเนินงานวิจัย

4.1 ผลการทดสอบแบบจำลอง

ผลจากการทดสอบรูปแบบโครงข่ายประสาทเทียม 10 รูปแบบ โดยกำหนดจำนวนรอบในการฝึกสอนเท่ากัน คือ 500 รอบ อัตราการเรียนรู้ เท่ากับ 0.1 และค่าโมเมนตัม เท่ากับ 0.2 จากนั้นทดลองหาจำนวนนิวรอนในชั้นซ่อนในรูปแบบ

ที่แตกต่างกัน เพื่อหาโครงข่ายประสาทเทียมที่ให้ค่าความคลาดเคลื่อนเฉลี่ยยกกำลังสองน้อยที่สุดดังตารางที่ 1

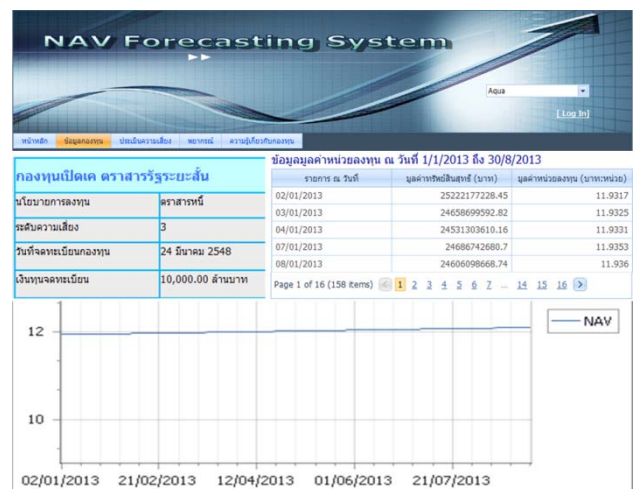
ตารางที่ 1 ผลการทดลองจากการสร้างแบบจำลอง

Model	Learning Rate	Momentum	MSE (Training)	MSE (Testing)
8-5-1	0.1	0.2	0.0024	0.0094
8-6-1	0.1	0.2	0.0019	0.0140
8-7-1	0.1	0.2	0.0017	0.0064
8-8-1	0.1	0.2	0.0016	0.0040
8-9-1	0.1	0.2	0.0015	0.0068
8-10-1	0.1	0.2	0.0011	0.0062
8-11-1	0.1	0.2	0.0016	0.0062
8-12-1	0.1	0.2	0.0013	0.0046
8-13-1	0.1	0.2	0.0014	0.0058
8-14-1	0.1	0.2	0.0013	0.0039

จากการทดลองพบว่ารูปแบบโครงข่ายประสาทเทียมที่เหมาะสมที่สุด คือ 8-10-1 ซึ่งมีจำนวนนิวรอนในชั้นนำเข้า 8 นิวรอน ชั้นซ่อน 10 นิวรอน และชั้นแสดงผล 1 นิวรอน และมีค่าความคลาดเคลื่อนเฉลี่ยยกกำลังสองน้อยที่สุดเท่ากับ 0.0062

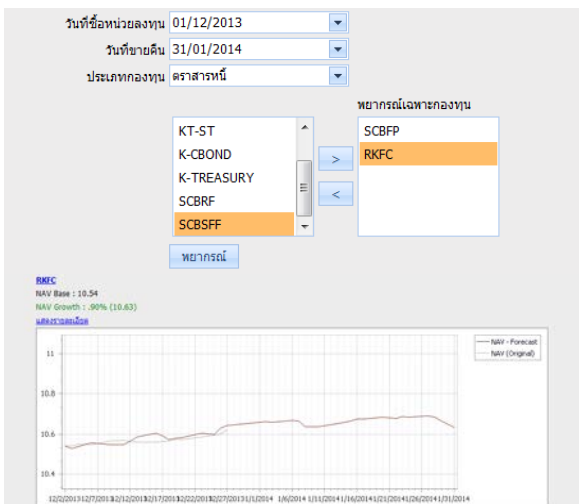
4.2 ตัวอย่างต้นแบบการพัฒนากระบวนการ

ตัวอย่างต้นแบบการพัฒนากระบวนการพยากรณ์ราคากองทุนรวม โดยใช้เหมืองข้อมูลและโครงข่ายประสาทเทียม สามารถเรียกดูราคากองทุนรวมย้อนหลังได้ตามช่วงเวลาที่ผู้ใช้งานต้องการ โดยระบบสามารถแสดงผลในรูปแบบกราฟพร้อมกับราคาเป็นรายวัน ดังภาพที่ 4



ภาพที่ 4: หน้าจอแสดงราคากองทุนรวมย้อนหลัง

ในส่วนการพยากรณ์ราคากองทุนรวมสามารถเลือกกองทุน และช่วงเวลาที่ต้องการพยากรณ์ได้ โดยระบบจะแสดงผลในรูปแบบกราฟพร้อมกับราคาเป็นรายวัน และสามารถแสดงผลการจัดอันดับตามอัตราผลตอบแทนที่พยากรณ์ได้ ดังภาพที่ 5



ภาพที่ 5: หน้าจอพยากรณ์ราคากองทุนรวม

5. สรุปผลการวิจัย และข้อเสนอแนะ

งานวิจัยนี้ได้นำหลักการ โครงข่ายประสาทเทียมแบบหลายชั้น ลักษณะการเรียนรู้แบบแพร่กระจายย้อนกลับมาสร้างแบบจำลองพยากรณ์ราคากองทุนรวมตราสารหนี้ โดยรวบรวมข้อมูลทางด้านการเงินและเศรษฐกิจ ตั้งแต่ปี พ.ศ. 2551 ถึง พ.ศ. 2556 และจากการทดลองทั้ง 10 รูปแบบพบว่าโครงข่ายประสาทเทียมที่เหมาะสมที่สุด คือ 8-10-1 โดยวัดจากความคลาดเคลื่อนเฉลี่ยกำลังสองน้อยที่สุด เท่ากับ 0.0062 ซึ่งสรุปได้ว่าแบบจำลองมีความเหมาะสม และสามารถนำไปประยุกต์ใช้งานได้จริง โดยในส่วนของการศึกษาครั้งต่อไปจะนำแบบจำลองไปพัฒนาระบบพยากรณ์ราคากองทุนรวมในรูปแบบของเว็บแอปพลิเคชัน เพื่อช่วยสนับสนุนการตัดสินใจลงทุนในกองทุนรวม และควรมุ่งเน้นในการเพิ่มความแม่นยำของการพยากรณ์ราคาให้มากขึ้น เช่น เพิ่มตัวแปรที่เป็นปัจจัยที่เกี่ยวข้องกับการพยากรณ์ราคา เป็นต้น

เอกสารอ้างอิง

[1] ศูนย์ส่งเสริมการพัฒนาความรู้ตลาดทุน สถาบันกองทุนเพื่อพัฒนาตลาดทุน. ตลาดการเงินและการลงทุนในหลักทรัพย์. พิมพ์ครั้งที่ 12. กรุงเทพฯ: อมรินทร์พริ้นติ้งแอนด์พับลิชชิ่ง, 2554

[2] ภรณ์ยา อัมฤครัตน์ . “การเปรียบเทียบประสิทธิภาพการคัดเลือกและจำแนกข้อมูลด้วยวิธีการทางเครือข่ายประสาทเทียม” in the 5TH National Conference on Computing and Information Technology, 2009, pp. 131-138.

[3] คีตา จาริก, “การพยากรณ์ราคาของพาราโดยใช้โครงข่ายประสาทเทียม”, กรุงเทพมหานคร: ปัญหาพิเศษ วิทยาศาสตร์มหาบัณฑิต ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ, มหาวิทยาลัยเทคโนโลยีพระจอมเกล้า พระนครเหนือ ,2555

[4] Amir Omidi, Esmaeil Nourani and Mehdi Jalili. “Forecasting stock prices using financial data mining and Neural Network”. *Computer Research and Development (ICCRD), 2011 3rd International Conference*; 2011 Mar 11-13; Shanghai, China.

[5] Yue Ma, Yu Chang and Chunyu Xia. “Applied Research on Stock Forecasting Model Based on BP neural network”. *International Conference on Electronic & Mechanical Engineering and Information Technology*; 2011 Aug 12-14; Harbin, China.

[6] D. Zhang, Q. Jiang, X. Li, “Application of Neural Networks in Financial Data Mining”, *International Journal of Computational Intelligence*, 2005

[7] พงษ์ศิริ ศิริพานิช ปรานี นิลกรณ์ และ สุดาตระการเถลิงศักดิ์. 2550. “การพยากรณ์อนุกรมเวลาด้วยตัวแบบผสม ARIMA และ เครือข่ายประสาทเทียม”. *การประชุมวิชาการด้านการวิจัยดำเนินงานแห่งชาติ ประจำปี 2550*. 6-7 กันยายน 2550

[8] Phaisarn Sutheebanjard and Wichian Premchaiswadi. “Stock Exchange of Thailand Index prediction using Back Propagation Neural Networks”. *Computer and Network Technology (ICCNT) 2010 Second International Conference*; 2010 April 23-25; Bangkok, Thailand.

[9] Rwei-Shan Lu, Shang-Wu Yu and Yi-Hsien Lin. “The Prediction of Applying Smooth Support Vector Regression and Back Propagation Network in Mutual Fund Performance”. *Neural Networks, 2008. IJCNN 2008. (IEEE World Congress on Computational Intelligence). IEEE International Joint Conference*; 2008 June 1-8; Hong Kong.

[10] จิรนนท์ วิโรจน์อึ้งยง และมาลีรัตน์ โสदानิล. “ระบบพยากรณ์ปริมาณการใช้หินปูนเพื่ออุตสาหกรรมซีเมนต์ โดยใช้โครงข่ายประสาทเทียม” in the 9TH National Conference on Computing and Information Technology, 2013, pp. 71-76.

พยากรณ์ลูกค้าที่มีความต้องการซื้อกองทุนรวมหุ้นระยะยาว (LTF)

โดยใช้เทคนิคการทำเหมืองข้อมูล

Forecasting Customer Demand for Long Term Equity Fund (LTF)

Using Data Mining Techniques

เชาวน์วัฒน์ ปรารณาแสงกุลศ (Chaowat Prattanasangkusol)¹ และ มาลีรัตน์ โสคันนิล (Maleerat Sodanil)²

¹ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

²ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹prattanasangkusol@hotmail.com, ²msn@kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อวิเคราะห์พฤติกรรม และพยากรณ์หาความต้องการของลูกค้าในการซื้อกองทุนรวมหุ้นระยะยาว (LTF) โดยใช้เทคนิคต้นไม้ตัดสินใจ เปรียบผลกับเทคนิคการวิเคราะห์การถดถอยโลจิสติกส์ เพื่อให้การพยากรณ์มีความถูกต้อง แม่นยำมากที่สุด ซึ่งเทคนิคที่ให้ค่าความถูกต้องสูงสุด คือ ต้นไม้ตัดสินใจ วิธี GINI แบบไม่ใช้ฟังก์ชันการคัดเลือกตัวแปร ซึ่งให้ค่าความถูกต้องร้อยละ 88.186 ผลการวิจัยดังกล่าวช่วยหาแบบแผน (Pattern) กลุ่มลูกค้าเป้าหมายที่คาดว่าจะสนใจซื้อกองทุนรวมหุ้นระยะยาว (LTF) และนำแบบจำลอง (Model) ที่ได้ช่วยในการค้นหาลูกค้า เพื่อนำเสนอโปรโมชั่น กระตุ้นยอดขาย ซึ่งเป็นการเพิ่มผลกำไรให้กับธุรกิจในอนาคตต่อไป

คำสำคัญ: ต้นไม้ตัดสินใจ การวิเคราะห์การถดถอยโลจิสติกส์ การพยากรณ์

Abstract

This research issue aims to analyze the customer behavior and to predict need of customer decision who purchase Long Term Equity Fund (LTF) by comparing the outcomes of Decision Tree and Logistic Regression. As the result, the best technique provided the highest accuracy is Decision Tree Approach - the non-selected variable function or GINI, given 88.186% accuracy. This research would provide the guidance to predict the target customers pattern expected to buy the Long Term Equity (LTF). Additionally, the model that helped to find

customers, offer promotion, increase sales and profits to the business in the future.

Keywords : Decision Tree, Logistic Regression, Forecast

1. บทนำ

คนส่วนใหญ่รู้จักแต่วิธีหาเงินให้ได้มากที่สุด ทุ่มเทเวลาศึกษาหาวิธีทำเงินเพิ่ม แต่ไม่รู้วิธีจัดการกับเงินที่หามาได้อย่างเหมาะสม การลงทุนเป็นวิธีการจัดการที่จะนำเงินที่หามาได้ไปทำประโยชน์ ให้ได้รับผลตอบแทนที่ดีขึ้นมากกว่าเพียงแค่การฝากเงินกับสถาบันการเงิน และโดยทั่วไปมักจะคิดว่าการลงทุนเป็นกระบวนการที่ซับซ้อนยุ่งยาก หรือต้องเป็นมืออาชีพเท่านั้น จึงจะสามารถลงทุนได้ แต่ความจริงแล้วการลงทุนไม่ยากเกินกว่าที่จะทำความเข้าใจ

สำหรับบุคคลทั่วไปที่มีความรู้ ความเข้าใจในด้านการลงทุนไม่มากนัก และสามารถยอมรับความเสี่ยงได้ การลงทุนในกองทุนรวมเป็นอีกหนึ่งทางเลือกที่น่าสนใจ เช่น กองทุนรวมตลาดเงิน, กองทุนรวมดัชนี, กองทุนรวมตราสารหนี้, กองทุนรวมตราสารทุน เป็นต้น ซึ่งแต่ละกองทุนจะมีความแตกต่างกันทั้งในเรื่องของนโยบายการลงทุน ผลตอบแทนที่ได้รับ ระดับความเสี่ยง และสิทธิประโยชน์ต่างๆ อย่างไรก็ตาม ยังมีกองทุนอีกรูปแบบหนึ่งซึ่งเป็นการออมในระยะยาว อีกทั้งยังได้รับสิทธิประโยชน์ทางด้านภาษี คือ กองทุนรวมหุ้นระยะยาว (LTF)

ดังนั้น การนำข้อมูลของลูกค้าที่เลือกลงทุนในกองทุนรวมต่างๆ มาทำการวิเคราะห์พฤติกรรม เพื่อพยากรณ์ความต้องการของลูกค้าที่ซื้อกองทุนรวมหุ้นระยะยาว (LTF) โดยใช้การทำ

เหมือนข้อมูล และนำแบบจำลอง (Model) ที่ช่วยในการค้นหากลุ่มลูกค้าเพื่อสร้างผลกำไรในอนาคตต่อไป

2. ทฤษฎีที่เกี่ยวข้อง

2.1 การทำเหมืองข้อมูล

การทำเหมืองข้อมูล [1][3] เป็นเทคโนโลยีใหม่ของการประยุกต์ใช้ข้อมูลที่เก็บอยู่ในฐานข้อมูลให้เกิดประโยชน์สูงสุดแก่หน่วยงานที่เป็นเจ้าของข้อมูล การประยุกต์ใช้ข้อมูลดังกล่าวถึงนี้มีได้หลายแนวทาง แต่โดยทั่วไปมักจะเป็นการสรุปภาพรวมของข้อมูลในฐานข้อมูล การวิเคราะห์แนวโน้มการเปลี่ยนแปลงของข้อมูล หรือ การค้นหาความสัมพันธ์ที่ซ่อนอยู่ภายในกลุ่มของข้อมูล

การประยุกต์ใช้เทคนิคการทำเหมืองข้อมูลจะมีได้หลากหลาย แต่สามารถจัดกลุ่มกว้าง ๆ ได้เป็นสองกลุ่ม ดังนี้

การทำเหมืองข้อมูลเพื่อการทำนาย

เป็นการนำความรู้ที่เรียนรู้มาจากข้อมูลที่มีอยู่เพื่อประโยชน์ในการทำนายข้อมูลใหม่ที่จะเกิดขึ้นในอนาคต เช่น จากข้อมูลลูกค้าของแผนกสินค้าของธนาคารที่ได้มีการจัดลำดับชั้นของลูกค้าไว้แล้วว่าใครเป็นลูกค้าชั้นดี ชั้นปานกลาง และลูกค้าที่มักจะผิดนัดชำระหนี้ โปรแกรมการทำเหมืองข้อมูลสามารถเรียนรู้ข้อมูลเหล่านี้ และค้นหาโมเดลที่สามารถใช้อธิบายลักษณะของลูกค้าชั้นดี ลูกค้าชั้นปานกลาง และลูกค้าที่ไม่เป็นที่ต้องการ จากโมเดลที่ได้สามารถนำไปใช้ทำนายลูกค้าใหม่ที่มาขอสินเชื่อได้ว่าเขาน่าจะเป็นลูกค้าประเภทใด

2.2 เทคนิคต้นไม้ตัดสินใจ (Decision Tree)

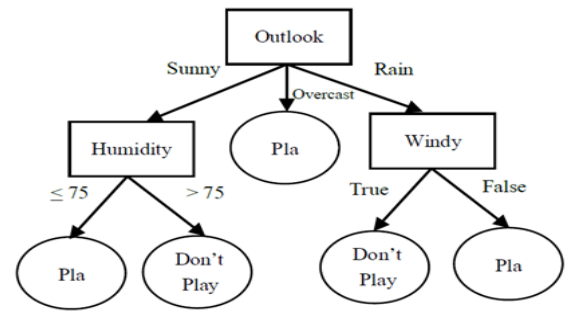
[1][3] เป็นเทคนิคที่ให้ผลลัพธ์ในลักษณะของโครงสร้างต้นไม้ ซึ่งเมื่อมีข้อมูลที่ต้องการจะจัดกลุ่มก็จะนำ attribute ต่าง ๆ ของข้อมูลนั้นไปเทียบกับ Decision tree ตามเส้นทางใน tree จนกระทั่งคลาสปลายทางซึ่งก็คือ กลุ่มของข้อมูลที่เหมือนกันโดยปกติมักประกอบด้วยกฎในรูปแบบ “ถ้า เงื่อนไข แล้ว ผลลัพธ์” เช่น

“If Income = High and Married = No THEN Risk = Poor”

“If Income = High and Married = Yes THEN Risk = Good”

โดยลักษณะของต้นไม้ตัดสินใจนี้มีลักษณะคล้ายกับต้นไม้กลับหัวโดยโหนดแรกสุดจะเป็นรากของต้นไม้ (Root node)

แต่ละโหนดแสดงคุณลักษณะ (attribute) แต่ละกิ่งจะแสดงค่าผลในการทดสอบ และ โหนดใบ (Leaf node) แสดงคลาสที่กำหนดไว้ ดังภาพที่ 1



ภาพที่ 1: ลักษณะของต้นไม้ตัดสินใจ

2.2.1 ขั้นตอนวิธี ID3 (ID3 Algorithm)

ขั้นตอนวิธี ID3 เป็นขั้นตอนวิธีที่ใช้ในการสร้างต้นไม้ตัดสินใจ โดยนำหลักการของทฤษฎีข่าวสารมาใช้ ค่าที่วัดได้จะนำมาตัดสินใจว่าจะใช้ตัวแปรใดในการแบ่งข้อมูลโดยวิธีการกำหนดโครงสร้างต้นไม้ตัดสินใจจะเป็นการเลือกข้อมูลตามลำดับของตัวชี้วัดหรือค่าเกน (Gain) สูงที่สุดเป็นข้อมูลเริ่มต้นและข้อมูลถัดไปที่มีค่าลดหลั่นกันตามลำดับ ตัวอย่างเช่น การพิจารณาจากกลุ่มข้อมูล 2 คลาส คือ P และ N โดยจำนวนตัวอย่างในคลาส P คือ p ตัว และจำนวนตัวอย่างในคลาส N คือ n ตัว ส่วนค่าของข้อมูลคือค่าคาดคะเนที่กลุ่มตัวอย่างต้องใช้จำนวนบิตในการแยกคลาส P และ N โดยนิยามดังสมการที่ 2-1

$$I(p, n) = -\frac{p}{p+n} \log_2 \left(\frac{p}{p+n} \right) - \frac{n}{p+n} \log_2 \left(\frac{n}{p+n} \right)$$

ค่าคาดคะเนของข้อมูล (Entropy) เป็นค่าที่แยกโดยการใส่ลักษณะประจำ A ซึ่งกำหนด A คือ ลักษณะประจำที่แบ่ง S ออกเป็น {S1, S2, ..., Sv} โดยให้ S1 มีตัวอย่างจากคลาส P จำนวน P1 และตัวอย่างจากคลาส N จำนวน n1 ดังสมการที่ 2-2

$$E(A) = \sum_{i=1}^v \frac{P_i + n_i}{p+n} I(p_i, n_i)$$

ดังนั้นค่าเกนข้อมูล (Data Gain) ที่ได้จากการแยกข้อมูลด้วยลักษณะประจำ A จะได้ดังสมการที่ 2-3

$$Gain(A) = I(p, n) - E(A)$$

2.2.2 ขั้นตอนวิธี C4.5 (C4.5 Algorithm)

ขั้นตอนวิธี C4.5 เป็นส่วนขยายของขั้นตอนวิธี ID3 ถูกพัฒนาโดย Ross Quinlan ใช้ในการสร้างต้นไม้ประกอบการตัดสินใจสำหรับใช้ในการจำแนกประเภทข้อมูล มีการใช้ค่าเกน

(Gain) และค่าคาดคะเนของข้อมูล (Entropy) เช่นเดียวกับ ID3 แต่จะมีส่วนที่เพิ่มเติมจากขั้นตอนวิธี ID3 ที่สำคัญดังนี้

2.2.2.1 สามารถใช้งานได้ทั้งข้อมูลที่มีคุณลักษณะแบบต่อเนื่อง (Continuous) และแบบไม่ต่อเนื่อง (Discrete) โดยในส่วนของข้อมูลแบบต่อเนื่องขั้นตอนวิธี C4.5 จะสร้างจุดเริ่ม (Threshold) และแยกคุณลักษณะนั้นออกเป็น 2 ส่วน คือ ส่วนที่มีค่ามากกว่าและน้อยกว่าและเท่ากับค่าที่ใช้ในการสร้างจุดเริ่ม

2.2.2.2 สามารถใช้กับข้อมูลฝึกสอน (Training Data) ที่ไม่มีค่าของคุณลักษณะได้ โดยจะทำการทำเครื่องหมายในคุณลักษณะนั้นเป็น “?” และไม่นำค่านั้นมาคำนวณหาค่าคาดคะเนของข้อมูล (Entropy)

2.2.2.3 สามารถใช้งานได้กับค่าที่มีความผิดปกติหรือมีความเสียหาย

2.2.2.4 สามารถทำการปรับแต่งต้นไม้ประกอบการตัดสินใจ (Pruning Trees) ในขณะที่สร้างได้

2.3 การวิเคราะห์การถดถอยโลจิสติกส์ (Logistic Regression Analysis)

[1][3][4] หลักการของการวิเคราะห์ความถดถอยแบบพหุคูณ ตัวแปรตาม (ตัวแปรเกณฑ์) จะเป็นตัวแปรเชิงปริมาณ ในขณะที่ตัวแปรอิสระ (ตัวแปรพยากรณ์) จะเป็นตัวแปรเชิงปริมาณเพียงอย่างเดียว หรืออาจมีตัวแปรบางตัวที่เป็นตัวแปรเชิงปริมาณและตัวแปรบางตัวเป็นตัวแปรเชิงกลุ่มหรือเชิงคุณภาพ แต่ถ้าตัวแปรตามเป็นตัวแปรเชิงกลุ่ม จะต้องใช้เทคนิคการวิเคราะห์การถดถอยโลจิสติกส์ ซึ่งยังคงมีวัตถุประสงค์และแนวคิดเหมือนกับการวิเคราะห์การถดถอยเชิงเส้น คือ เพื่อศึกษาความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระ และนำเสนอการที่ได้ไปประมาณหรือพยากรณ์ตัวแปรตาม เมื่อกำหนดค่าตัวแปรอิสระ

2.4 การคัดเลือกคุณลักษณะ (Feature Selection)

[1][3] การคัดเลือกคุณลักษณะ เป็นส่วนการทำงานที่เลือกคุณลักษณะเด่นที่ได้จากการสร้างและค้นหาคุณลักษณะเด่นเพื่อหาคุณลักษณะที่เหมาะสมที่สุดสำหรับแต่ละงาน ซึ่งส่วนใหญ่จะเป็นการหาจำนวนลักษณะที่น้อยที่สุดเพื่อให้ความซับซ้อนของการคำนวณน้อย แต่ให้ผลการจำแนกประเภทข้อมูลได้ผลดีที่สุด โดยวิธีที่เลือกใช้เพื่อเปรียบเทียบผลของค่า เทคนิคการคัดเลือกข้อมูลในการทำเหมืองข้อมูลมีวัตถุประสงค์หลักสามประการ คือ เพื่อเพิ่มประสิทธิภาพการทำนาย เพื่อการ

สังเคราะห์โมเดลได้อย่างรวดเร็ว และเพื่อลดความซับซ้อนของรูปแบบโมเดล

Feature Selection คือการเลือกเซตย่อยของคุณลักษณะที่สามารถอธิบาย Output Y โดยที่ Y หมายถึงตัวแปรที่ต้องการเรียน อัลกอริทึมในการเลือกคุณลักษณะสามารถจัดได้เป็น 3 ประเภทใหญ่ ขึ้นอยู่กับว่าอัลกอริทึมนั้น ๆ เกี่ยวข้องกับการทำนายที่จะนำคุณลักษณะที่เลือกได้ไปเรียนรู้ต่อ เช่น Support Vector Machine, linear regression, Naive Bayes เป็นต้น

2.5 การวัดประสิทธิภาพของการพยากรณ์ด้วย ROC Curve

[2] ประสิทธิภาพของการพยากรณ์ อาจวัดด้วยค่า Sensitivity และ Specificity ภายใต้ตัวอย่างขนาด n โดยพิจารณาผลของการพยากรณ์ด้วยจำนวนเหตุการณ์ ดังตารางที่ 1

ตารางที่ 1: Confusion Matrix จากผลของการพยากรณ์

เหตุการณ์	ผลของการพยากรณ์		รวม
	ถูก ($\hat{y} = 1$)	ไม่ถูก ($\hat{y} = 0$)	
สนใจ ($y = 1$)	a	b	a+b
ไม่สนใจ ($y = 0$)	c	d	c+d
รวม	a+c	b+d	a+b+c+d

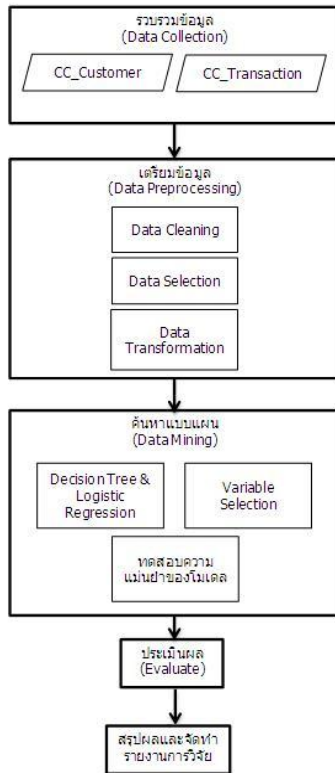
2.6 งานวิจัยที่เกี่ยวข้อง

อุโฆษ (2550) การพัฒนาระบบที่สนับสนุนการตัดสินใจมาช่วยจัดกลุ่มข้อมูลซึ่งจะช่วยในการตัดสินใจของนักลงทุน ว่าควรลงทุนในกองทุนใดบ้าง โดยได้ทำการประเมินคุณภาพของการวิเคราะห์ข้อมูลกองทุนรวม จากการสำรวจพฤติกรรมการลงทุนของผู้ใช้ระบบมาวิเคราะห์ด้วยการจัดกลุ่มกองทุนที่มีลักษณะเหมือนกันในทุกปัจจัยแล้วใช้แผนผังต้นไม้ตัดสินใจช่วยหากองทุนรวมที่เหมาะสมต่อนักลงทุน

3. วิธีดำเนินการวิจัย

3.1 แผนผังขั้นตอนการดำเนินงาน

การรวบรวมข้อมูล (Data Collection) จาก Enterprise Data Warehouse และฐานข้อมูลจากระบบกองทุนในรูปแบบ SAS Dataset ด้วยโปรแกรม SAS 9.3 โดยแบ่งออกเป็น 2 ตาราง (Dataset) ดังนี้



ภาพที่ 2: แผนผังขั้นตอนการดำเนินงาน

3.2 การรวบรวมข้อมูล (Data Collection)

รวบรวมข้อมูล (Data Collection) จาก Enterprise Data Warehouse และฐานข้อมูลจากระบบกองทุนในรูปแบบ SAS Dataset ด้วยโปรแกรม SAS 9.3 โดยแบ่งออกเป็น ตารางที่ 2: จำนวนแอททริบิวต์ทั้งหมดที่นำมาสร้างโมเดล

Name	Description	Role	Level
ID	ID	Input	Nominal
AGE	อายุ	Input	Interval
Annual_Premium	ค่าเบี้ยประกันชีวิต (รายปี)	Input	Interval
CAR_OWNERSHIP	จำนวนรถที่มี (คัน)	Input	Nominal
CHILDREN	มีบุตร	Input	Nominal
F_LTF	ถ้า > 0 คือ มีกองทุนLTF	Rejected	Interval
F_RMF	ถ้า > 0 คือ มีกองทุนRMF	Input	Interval
F_TMF	ถ้า > 0 คือ มีกองทุนTMF	Input	Interval
F_TF	ถ้า > 0 คือ มีกองทุน Tern Fund	Input	Interval
F_OEF_OTH	ถ้า > 0 คือ มีกองทุนอื่นๆ	Input	Interval
F_SFF	ถ้า > 0 คือ มีกองทุนSFF	Input	Interval
HOME_OWNERSHIP	จำนวนบ้านที่มี (หลัง)	Input	Nominal
OCC_Group	อาชีพ	Input	Nominal
Proxy_Income	รายได้	Input	Interval
SEX	เพศ	Input	Nominal
Wealth_avg	สินทรัพย์เฉลี่ย 6 เดือน	Input	Interval
TARGET_LTF	TARGET_LTF	Target	Binary

3.3 การเตรียมข้อมูล (Data Preprocessing)

3.3.1 การทำข้อมูลให้สมบูรณ์ (Data Cleansing) ปรับเปลี่ยนข้อมูลบางคอลัมน์เพื่อให้สามารถนำไปใช้ในการทำเหมืองข้อมูลได้

3.3.1.1 ข้อมูลไม่สมบูรณ์ (incomplete data) เช่น ค่าของคุณลักษณะขาดหาย (missing value) ขาดคุณลักษณะที่น่าสนใจหรือขาดรายละเอียดของข้อมูล

3.3.1.2 ข้อมูลรบกวน (noisy data) เช่น ข้อมูลมีค่าผิดพลาด (error) หรือมีค่าผิดปกติ (Outliers)

3.3.1.3 ข้อมูลไม่สอดคล้อง (Inconsistent data) เช่น ข้อมูลเดียวกัน แต่ตั้งชื่อต่างกัน หรือใช้ค่าแทนข้อมูลที่ต่างกัน

3.3.2 การคัดเลือกข้อมูล (Data Selection) ให้มีความเหมาะสมที่คาดว่าจะมีผลกระทบต่อกระบวนการวิเคราะห์ตรงประเด็น เพื่อให้ได้ข้อมูลที่มีคุณภาพสำหรับการพิจารณา

3.3.3 การปรับเปลี่ยนรูปแบบข้อมูล (Data Transformation) ให้อยู่ในรูปแบบเดียวกันที่เป็นมาตรฐานและเหมาะสมที่จะนำไปใช้กับแบบจำลอง (Model)

3.4 การค้นหาแบบแผน (Pattern)

3.4.1 ใช้เทคนิค Decision Tree วิธี GINI (ใช้อัลกอริทึม CART) และวิธี Entropy (ใช้อัลกอริทึม C4.5) และเทคนิค Logistic Regression Analysis เพื่อเปรียบเทียบหาเทคนิคและโมเดลที่สามารถพยากรณ์ได้แม่นยำและมีประสิทธิภาพสูงที่สุด โดยวัดจากค่าความถูกต้อง (Accuracy) ของแต่ละโมเดล

3.4.2 ใช้ฟังก์ชัน Variable Selection เพื่อหาขนาดมิติของข้อมูลที่มีความเหมาะสม เพื่อให้การพยากรณ์ยังคงความแม่นยำอยู่และช่วยเพิ่มประสิทธิภาพในการทำแบบจำลอง (Model)

3.4.3 แบ่งข้อมูลทั้งหมดจาก 100 % ออกเป็นข้อมูลสำหรับ Training Set อัตราส่วน 45 % เพื่อหาหน้าหนักความสัมพันธ์ของตัวแปรแต่ละตัว และตรวจสอบความถูกต้อง (ความเชื่อมั่น) ด้วย Validation Set อัตราส่วน 35% ว่าจะสามารถพยากรณ์ได้โมเดลที่ดีที่สุด และทดสอบโมเดลที่ได้กับข้อมูลที่เราต้องการพยากรณ์ Testing Set อัตราส่วน 20% ตามลำดับ

4. ผลการดำเนินงาน

4.1 ผลการพยากรณ์

ใช้ข้อมูลการทำรายการลูกค้าของธนาคารพาณิชย์แห่งหนึ่งที่มีการซื้อกองทุนตั้งแต่ปี 2554 จนถึงปี 2556 จำนวนข้อมูลที่ใช้สำหรับสร้างโมเดลประกอบด้วย

F_LTF = 1 คือ ลูกค้าที่ซื้อกองทุนรวมหุ้นระยะยาว (LTF)

F_LTF = 0 คือ ลูกค้าไม่ซื้อกองทุนรวมหุ้นระยะยาว (LTF)

ตารางที่ 3: จำนวนข้อมูลทั้งหมดสำหรับทำโมเดล

Target	#Customers	%
0	419,738	77.41%
1	122,492	22.59%
Total	542,230	100.00%

โดยแบ่งข้อมูลสำหรับ Training 45% และข้อมูลสำหรับ Validation 35% ดังตารางที่ 4

ตารางที่ 4: จำนวนข้อมูลสำหรับ Train และ Validation

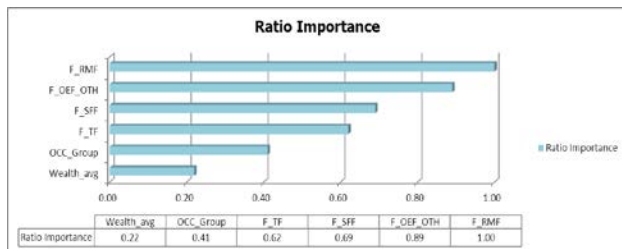
F_LTF	#Train	%Train	#Validation	%Validation
0 (ลูกค้าไม่ซื้อ LTF)	188,881	77.41%	146,908	77.41%
1 (ลูกค้าซื้อ LTF)	55,121	22.59%	42,871	22.59%
Total	244,002	100.00%	189,779	100.00%

โมเดลที่ให้ค่าความถูกต้องมากที่สุด คือ Decision Tree วิธี GINI แบบไม่ใช้ฟังก์ชัน Variable Selection

ตารางที่ 5: โมเดล Decision Tree วิธี GINI แบบไม่ใช้ฟังก์ชัน Variable Selection

Data Role	FALSE		TRUE	
	Negative	Negative	Positive	Positive
TRAIN	16,697	176,762	12,119	38,424
VALIDATE	12,997	137,484	9,424	29,874

ข้อมูลที่ใช้ Training ให้ค่าความถูกต้อง (Accuracy) 88.19% และข้อมูลที่ใช้ Validation ให้ค่าความถูกต้อง (Accuracy) 88.19% โมเดลได้กำหนดค่าความสำคัญของตัวแปรที่มีผลกับโมเดล ดังภาพที่ 3

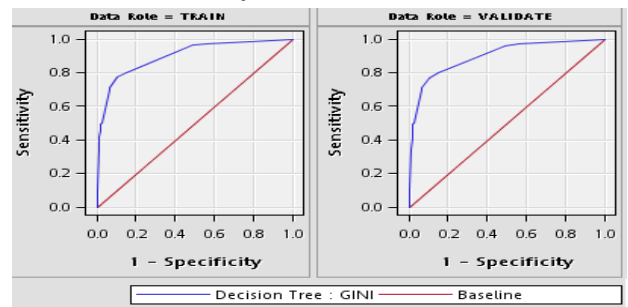


ภาพที่ 3: กราฟแสดงค่าความสำคัญของตัวแปรสำหรับโมเดล Decision Tree วิธี GINI แบบไม่ใช้ฟังก์ชัน Variable Selection

จากภาพที่ 3 พบว่า ตัวแปรที่มีความสำคัญกับโมเดลจากมากไปน้อย คือ กองทุน RMF (F_RMF), กองทุนอื่นๆ

(F_OEF_OTH), กองทุน SFF (F_SFF), กองทุน Term Fund (F_TF), อาชีพ (OCC_Group), สินทรัพย์เฉลี่ย 6 เดือน (Wealth_avg) ตามลำดับ

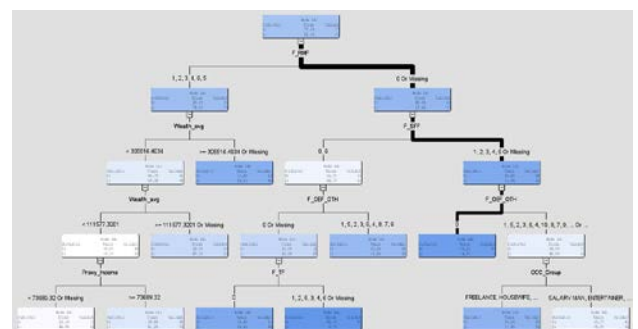
ผลของการพยากรณ์สามารถแสดงได้โดย ROC Curve ซึ่งเป็นวิธีแสดงอำนาจการจำแนกที่สำคัญในการพยากรณ์ ณ จุดตัด (Cut point) ของตัวแปรตอบสนอง ถ้าเปลี่ยนจุดตัดมีผลให้ค่า Sensitivity และ Specificity จะเปลี่ยนไปด้วย จุดตัดที่ดีที่สุดควรอยู่ที่จุดวกกลับของเส้นโค้ง ซึ่งยิ่งชันยิ่งดี เพราะจะให้พื้นที่ใต้เส้นโค้งด้านซ้ายของเส้นทแยงมุมหรือการพยากรณ์ถูกต้องเพิ่มขึ้นนอกจากนี้ การใช้ ROC Curve ยังเป็นวิธีการที่ดีในการวัดอำนาจการพยากรณ์ของตัว เมื่อเพิ่มตัวแปรอธิบายที่มีความสำคัญเข้าในตัวแบบ เส้นโค้งจะชันมากขึ้นและมีพื้นที่การพยากรณ์ได้ถูกต้องใต้เส้นโค้งมากขึ้นด้วย



ภาพที่ 4: ROC Curve ของโมเดล Decision Tree วิธี GINI แบบไม่ใช้ฟังก์ชัน Variable Selection

ตัวแปรลักษณะแผนผังต้นไม้ตัดสินใจที่ได้จากโมเดล Decision Tree วิธี GINI ใช้ฟังก์ชัน Variable Selection แสดงตัวแปรที่มีผลกับโมเดลและเงื่อนไขเพื่อใช้ในการพยากรณ์หาลูกค้าที่มีความต้องการซื้อกองทุนรวมหุ้นระยะยาว (LTF) โดยการเลือกช่องทางที่มีเปอร์เซ็นต์การทำนาย Target เป็น 1 (ซื้อกองทุนรวมหุ้นระยะยาว) สูงที่สุด

แผนผังต้นไม้ตัดสินใจที่ได้จากโมเดล Decision Tree วิธี GINI แบบไม่ใช้ฟังก์ชัน Variable Selection ดังนี้



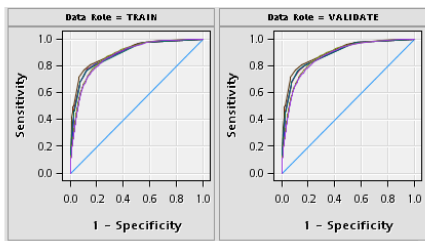
ภาพที่ 5: แผนผังต้นไม้ตัดสินใจที่ได้จากโมเดล Decision Tree วิธี GINI แบบไม่ใช้ฟังก์ชัน Variable Selection

ตารางที่ 6: การเปรียบเทียบค่าความถูกต้องของทั้ง 6 โมเดล

(Validation)

Model	Variable Selection	FALSE Negative	TRUE Negative	FALSE Positive	TRUE Positive	Accuracy
GINI	N	: 2,997	137,434	9,424	29,874	88.186%
Entropy		: 4,756	136,645	10,263	28,115	86.82%
Regression		: 21,815	138,812	3,096	21,056	84.24%
GINI(2)	Y	: 4,008	135,969	10,939	28,863	86.85%
Entropy (2)		: 4,008	135,969	10,939	28,863	86.85%
Regression (2)		: 22,748	138,527	3,381	20,123	83.60%

จากตารางที่ 6 เปรียบเทียบค่าความถูกต้อง (Accuracy) ที่ได้จากทั้ง 6 โมเดล จะเห็นได้ว่า ค่าความถูกต้อง (Accuracy) จากโมเดล Decision tree วิธี GINI แบบไม่ใช้ฟังก์ชัน Variable Selection ให้ค่าความถูกต้องสูงที่สุดทั้งข้อมูลที่ใช้ Validation เท่ากับ 88.186%



ภาพที่ 6: เปรียบเทียบ ROC Curve ของทั้ง 6 โมเดล

5. สรุปผลการวิจัยและข้อเสนอแนะ

5.1 การวิเคราะห์ข้อมูลโดยใช้ Decision Tree วิธี GINI แบบไม่ใช้ฟังก์ชัน Variable Selection นั้นทำให้การพยากรณ์มีความถูกต้อง (Accuracy) สูงถึง 88.186 % ซึ่งแสดงให้เห็นว่าโมเดลที่ได้นั้นมีประสิทธิภาพ

5.2 จากผลการศึกษาพฤติกรรมการซื้อกองทุนของลูกค้าพบว่า ลูกค้าที่มีกองทุนสำรองเลี้ยงชีพ (RMF), กองทุนรวมตราสารหนี้, กองทุนรวมอื่นๆ จะนำเงินมาซื้อกองทุนรวมหุ้นระยะยาว (LTF) ตามลำดับ

ข้อเสนอแนะ

เนื่องจากข้อมูลที่เป็นตัวแปรในการสร้างแบบจำลองในการทดลองนี้ไม่สามารถที่จะนำมาทำเหมืองข้อมูลเพื่อการอธิบายหาความสัมพันธ์ข้อมูลที่เชื่อมโยงกันได้ ดังนั้นในด้านข้อมูลควรที่จะเพิ่ม ตัวแปรในการสร้างแบบจำลองเพื่อใช้ค้นหารูปแบบที่น่าสนใจจากกลุ่มข้อมูลดังกล่าว เพื่อต่อยอดนำเสนอผลิตภัณฑ์กองทุนอื่นๆ ให้เหมาะสมกับพฤติกรรมของลูกค้าแต่ละคนได้

เอกสารอ้างอิง

- [1] กฤษณะ ไวยมัย, ชิดชนก สงศิริ และชนวินทร์ ภัทธรมา นนท์. "การใช้เทคนิคคาด้าไมน์นิ่งเพื่อพัฒนาคุณภาพการศึกษานิสิตคณะวิศวกรรมศาสตร์." *The Nectec Technical Journal*.2001.
- [2] กันยารพ หาญกล้า. "การเปรียบเทียบตัวแบบลอจิสต์และการใช้ ROC Curve ในการวิเคราะห์ปัจจัยที่ส่งผลต่อระดับค่าจ้าง." *วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาสถิติ บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร*, 2555.
- [3] พยูน พาณิชย์กุล. "การพัฒนาแบบคาด้าไมน์นิ่งโดยใช้ Decision Tree." *โครงการพัฒนาระบบงานปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ แขนงวิชาวิทยาการสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าคุณทหารลาดกระบัง*, 2548.
- [4] สวัสดิ์ชัย ศรีพนมชนากร. "การวิเคราะห์การถดถอยโลจิสติก." (ออนไลน์). สืบค้นจาก <http://www.thairenu.com/logistic.htm>
- [5] กัลยา วานิชย์บัญชา. "การวิเคราะห์ข้อมูลหลายตัวแปร." *พิมพ์ครั้งที่ 3. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย*, 2551.
- [6] ชนียงศักดิ์ กิริตวานิชย์ และ ภัสรา ขวาลกร. "รู้วิเคราะห์เจาะเรื่องกองทุนรวม." *กรุงเทพมหานคร : บุญศิริการพิมพ์ จำกัด*, 2551.
- [7] เพชร ชุมทรัพย์. "หลักการลงทุน." *พิมพ์ครั้งที่ 12. กรุงเทพมหานคร : มหาวิทยาลัยธรรมศาสตร์*, 2544.
- [8] จิตารีย์ รกรากทอง. "ปัจจัยที่มีผลต่อการตัดสินใจลงทุนในกองทุนรวมหุ้นระยะยาว" *สารนิพนธ์ปริญญาเศรษฐศาสตรมหาบัณฑิต คณะเศรษฐศาสตร์ มหาวิทยาลัยรามคำแหง*, 2549.

การวิเคราะห์ภาวะทางการเงินของสหกรณ์โดยประยุกต์ใช้ วิธีการวิเคราะห์ค่าเฉลี่ยและแสดงผลในรูปแบบธุรกิจอัจฉริยะ Cooperative Financial Analysis with CAMELS Analysis and Business Intelligence

พรชัยพล พรหมมาศ (Patsapol Prommat)¹ และผุสดี บุญรอด (Pudsadee Boonrawd)²

^{1,2} ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
patsapol@dmf.go.th¹, pudsadee@kmutnb.ac.th²

บทคัดย่อ

งานวิจัยนี้เสนอการสร้างแบบจำลองการวิเคราะห์ภาวะทางการเงินของสหกรณ์ โดยประยุกต์ใช้วิธีการวิเคราะห์ค่าเฉลี่ย ซึ่งเป็นตัววัดสถานภาพทางการเงินที่สถาบันเงินต่าง ๆ ให้การยอมรับ โดยแบ่งการวิเคราะห์ออกเป็น 6 มิติ คือ ความเพียงพอของเงินทุนต่อความเสี่ยง คุณภาพของสินทรัพย์ จีดีความสามารถในการบริหาร การทำกำไร สภาพคล่อง และผลกระทบต่อธุรกิจ ร่วมกับการหาความสัมพันธ์และการแบ่งประเภทข้อมูล จากข้อมูลทางการเงินของสหกรณ์ ตั้งแต่ปี พ.ศ. 2552 - 2556 เพื่อให้สหกรณ์สามารถรู้ถึงสถานภาพทางการเงินของตนเอง และยังสามารถเทียบเคียงกับสหกรณ์อื่น ๆ ผ่านข้อมูลและอัตราส่วนสำคัญของสหกรณ์ ผลการวิจัยพบว่า กฎความสัมพันธ์ที่ได้มีค่าความเชื่อมั่นมากกว่า 0.95 และจากการทดสอบความถูกต้องของแบบจำลองแบบไขว้ 10 ส่วนพบว่าแบบจำลองการวิเคราะห์ภาวะทางการเงินที่สร้างขึ้นมีความแม่นยำถึง 98.71% ดังนั้นสามารถสรุปได้ว่าการวิเคราะห์ภาวะทางการเงินของสหกรณ์โดยประยุกต์ใช้วิธีการวิเคราะห์ค่าเฉลี่ยสามารถวิเคราะห์ภาวะทางการเงินของสหกรณ์ได้อย่างถูกต้อง และมีประสิทธิภาพ

คำสำคัญ: วิธีการวิเคราะห์ค่าเฉลี่ย ภาวะทางการเงิน เหมืองข้อมูล การแบ่งประเภทข้อมูล กฎความสัมพันธ์

Abstract

This research presents the development of a cooperative financial analysis model using CAMELS analysis in 6 dimension such as Capital Strength, Asset Quality, Management Ability, Earning Sufficiency,

Liquidity and Sensitivity with Association Rules and Classification and the data set was used to create model from cooperative financial data during 2008-2013 to help cooperative CEO's evaluate their financial performance. Moreover, it also provides the ability compare results to other cooperative in the same group though Peer Groups. According to the evaluation results, confident equaled 0.95 with accuracy of the model at 98.71% by 10-folds cross validation. In conclusion, the cooperative financial analysis model with CAMELS analysis achieved all research goals accurately and efficiently.

Keywords: CAMELS Analysis, Financial analysis, Data Mining, Classification, Association Rules

1. บทนำ

วิกฤตเศรษฐกิจในรอบหลายปีที่ผ่านมาเป็นเครื่องพิสูจน์ว่าสถาบันการเงินมีความสำคัญต่อระบบเศรษฐกิจของประเทศ เช่น วิกฤตแฮมเบอร์เกอร์ ซึ่งเกิดจากการที่สถาบันการเงินปล่อยสินเชื่อด้านอสังหาริมทรัพย์และสินค้าอื่น ๆ ให้กับผู้กู้โดยไม่จำกัดขอบเขต ทำให้มีลูกหนี้ด้อยคุณภาพจำนวนมาก วิกฤตเศรษฐกิจในครั้งนี้ที่ความรุนแรงจนส่งผลกระทบต่อประเทศสหรัฐอเมริกาและยังขยายไปทั่วโลก อีกทั้งยังส่งผลกระทบต่อเนื่องมาจนถึงปัจจุบัน [1] สหกรณ์เป็นองค์กรธุรกิจรูปแบบหนึ่ง มีการบริหารและการดำเนินงานเหมือนสถาบันการเงินทั่วไป การมีสถานภาพทางการเงินที่มั่นคงจะส่งผลให้สหกรณ์สามารถดำเนินงานได้อย่างมีประสิทธิภาพ และเป็นเกราะป้องกันความเสี่ยงในด้านต่าง ๆ ที่จะเกิดขึ้น

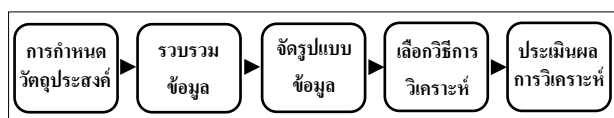
กับสหกรณ์ รวมทั้งสามารถสะท้อนภาพการดำเนินงานของสถาบันการเงินภาคประชาชนไปสู่แนวทางการพัฒนาสถาบันการเงินของประเทศ อันจะเป็นประโยชน์และส่งผลดีต่อสภาพสังคมและเศรษฐกิจของประเทศในอนาคต [2] แต่กระบวนการในการวิเคราะห์และวัดผลทางการเงินนั้น ต้องอาศัยบุคลากรที่มีความรู้ ความชำนาญในหลาย ๆ แขนงวิชา เช่น ด้านสถิติ ด้านการบริหาร การเงินการบัญชี และต้องมีความเข้าใจในโครงสร้างงบการเงินของสหกรณ์ซึ่งเป็นเรื่องยากสำหรับสหกรณ์ที่จะมีบุคลากรที่มีความรู้ ความชำนาญรอบด้านดังกล่าว

คณะผู้วิจัยจึงได้เสนอแบบจำลองการวิเคราะห์ภาวะทางการเงินของสหกรณ์โดยนำเอาวิธีการวิเคราะห์คามेल 6 มิติ ซึ่งเป็นวิธีการวัดสถานภาพทางการเงินและประสิทธิภาพการดำเนินงานที่เป็นมาตรฐานและเชื่อถือได้มาประยุกต์ใช้ในการวิเคราะห์งบการเงินของสหกรณ์ เพื่อให้ผู้บริหารสหกรณ์เห็นถึงสถานภาพทางการเงินของสหกรณ์ของตน ช่วยให้เกิดการเฝ้าระวังทางการเงิน และมุ่งเน้นให้สหกรณ์สามารถประเมินตนเองให้รู้จักตนเองและสามารถเปรียบเทียบกับสหกรณ์ของตนกับสหกรณ์อื่น ๆ ในกลุ่มเดียวกันเพื่อเป็นแรงผลักดันให้สหกรณ์พัฒนาตนเองอย่างต่อเนื่อง ซึ่งเนื้อหาของงานวิจัยแบ่งออกเป็น 5 หัวข้อหลัก คือ บทนำ ทฤษฎีและงานวิจัยที่เกี่ยวข้อง วิธีการดำเนินการวิจัย ผลการดำเนินการวิจัย สรุปผลการวิจัยและข้อเสนอแนะ

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การวิเคราะห์งบการเงิน

การวิเคราะห์งบการเงิน หมายถึง กระบวนการประเมินผลการดำเนินงานของกิจการใดกิจการหนึ่งจากงบการเงิน เพื่อค้นหาข้อเท็จจริงเกี่ยวกับฐานะทางการเงินและผลการดำเนินงานของกิจการพร้อมทั้งนำข้อเท็จจริงดังกล่าวมาประกอบการตัดสินใจ [1] ซึ่งงบการเงินประกอบด้วย งบดุล งบกำไรขาดทุน งบกระแสเงินสด และหมายเหตุประกอบงบการเงิน โดยขั้นตอนการวิเคราะห์งบการเงินแสดงดังภาพที่ 1



ภาพที่ 1: ขั้นตอนการวิเคราะห์งบการเงิน

จากภาพที่ 1 ขั้นตอนการวิเคราะห์งบการเงินเริ่มจากกำหนดวัตถุประสงค์ในการวิเคราะห์งบการเงิน ทำการรวบรวมข้อมูลงบการเงินของกิจการที่ต้องการวิเคราะห์และจัดรูปแบบข้อมูลให้อยู่ในรูปแบบที่เหมาะสมในการนำไปใช้วิเคราะห์ จากนั้นทำการเลือกวิธีการวิเคราะห์งบการเงินซึ่งมีด้วยกันหลายแบบ เช่น การวิเคราะห์ตามแนวดิ่ง ตามแนวนอน หรือการวิเคราะห์อัตราส่วนทางการเงิน และนำผลที่ได้จากการวิเคราะห์งบการเงินมาแปลความหมายพร้อมทั้งประเมินผลการวิเคราะห์

ปัจจุบันมีวิธีการวิเคราะห์งบการเงินที่นิยมนำมาใช้เป็นเครื่องมือวัดผลทางการเงินของสถาบันการเงินคือ วิธีการ PATROL เป็นเครื่องมือที่ธนาคารในประเทศอิตาลีนำมาใช้เพื่อวัดสถานะทางการเงินของตนเอง วิธีการ ORAP เป็นวิธีการที่ถูกนำมาใช้เพื่อตรวจสอบความเสี่ยงที่อาจจะเกิดขึ้นกับธนาคารในประเทศฝรั่งเศสโดยวิเคราะห์จากองค์ประกอบทั้งหมด 14 องค์ประกอบ วิธีการวิเคราะห์คามेल (CAMELS Analysis) เป็นวิธีการที่ใช้วัดภาวะทางการเงินและประสิทธิภาพในการบริหารที่สถาบันการเงินในสหรัฐนิยมนำมาใช้ โดยแบ่งการวิเคราะห์ออกเป็น 6 มิติหรือวิธีการวิเคราะห์เพิร์ล (PEARLS Analysis) เป็นวิธีการที่ใช้วัดสถานภาพทางการเงินของกลุ่มเครดิตยูเนียน [3] แต่วิธีการวิเคราะห์คามेलให้ความเข้าใจและเหมาะสมที่จะนำมาใช้กับสหกรณ์ เพราะมีความอิสระคล่องตัวในการนำอัตราส่วนทางการเงินมาปรับใช้ได้อย่างหลากหลายผ่านมิติทั้ง 6 มิติ [2] โดยผลการวิเคราะห์คามेलสามารถนำไปเทียบเคียงกับสหกรณ์ที่มีขนาด ประเภท ช่วงเวลาและสถานการณ์เดียวกัน ผ่านข้อมูลและอัตราส่วนสำคัญของสหกรณ์เฉลี่ย (Peer Groups) ซึ่งเป็นค่าเฉลี่ยของข้อมูลและอัตราส่วนของสหกรณ์เฉพาะที่ดำเนินงาน และไม่มีส่วนขาดแห่งทุน [4] เพื่อเป็นแรงผลักดันให้เกิดการพัฒนาตนเองได้

2.2 เหมืองข้อมูล (Data Mining)

การทำเหมืองข้อมูล เป็นการคัดกรองสารสนเทศจากข้อมูลขนาดใหญ่ เพื่อค้นหารูปแบบและความสัมพันธ์ที่ซ่อนอยู่โดยอาศัยหลักสถิติ การรู้จำการเรียงตัวของเครื่องและหลักคณิตศาสตร์ ซึ่งมีรูปแบบในการคัดกรองหลายรูปแบบ เช่น การจำแนกประเภทข้อมูล (Classification) การแบ่งกลุ่มข้อมูล (Clustering) การหากฎความสัมพันธ์ (Association rules)

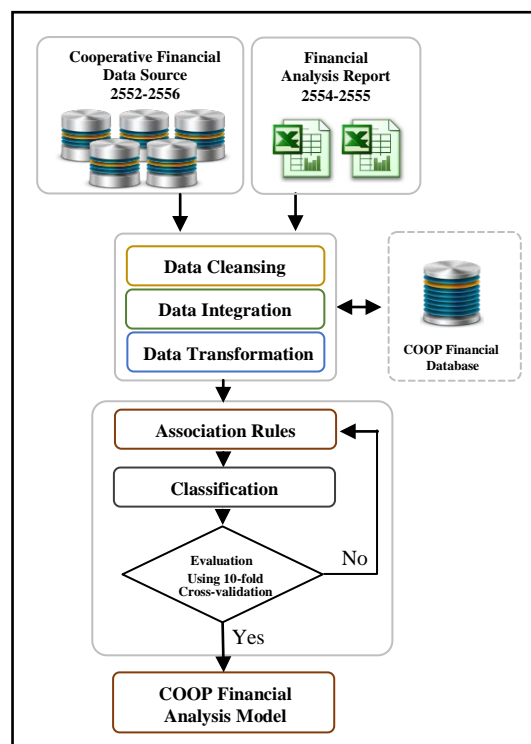
เป็นต้น งานวิจัยนี้เลือกใช้วิธีการแบ่งประเภทของข้อมูลแบบต้นไม้ตัดสินใจ (Decision Tree) ซึ่งเป็นการจำแนกประเภทข้อมูลซึ่งมีลักษณะเหมือนโครงสร้างต้นไม้ มีโหนดที่แสดงคุณลักษณะของข้อมูล (Attribute) ซึ่งแต่ละโหนดจะมีกิ่งเพื่อแสดงผลในการทดสอบของข้อมูล และลิฟโหนด (Leaf Node) แสดงถึงประเภทของข้อมูลที่กำหนดไว้ ซึ่งต้นไม้ตัดสินใจสามารถสร้างได้ง่าย และเลือกเฉพาะคุณลักษณะของข้อมูล (Attribute) ที่สำคัญในการสร้างแบบจำลอง [6], [7] และใช้วิธีการค้นหาความสัมพันธ์ เพื่อค้นหาความสัมพันธ์ของข้อมูลจากฐานข้อมูลที่มีขนาดใหญ่ที่มีอยู่ โดยมีค่าสนับสนุน (Support) และค่าความเชื่อมั่น (Confidence) เป็นค่าที่ใช้สำหรับบอกความน่าเชื่อถือของกฎ [8] วิธีการที่นิยมใช้ในการหาความสัมพันธ์คือ วิธีการเอปไรอริ (Apriori) ซึ่งมีจุดเด่นอยู่ที่ความเร็วในการค้นหาไอเท็มเซตที่เกิดขึ้นบ่อย ปัจจุบันการทำเหมืองข้อมูลได้ถูกนำไปประยุกต์ใช้ในงานหลายประเภททั้งในด้านธุรกิจที่ช่วยในการตัดสินใจของผู้บริหารในด้านวิทยาศาสตร์และการแพทย์รวมทั้งในด้านเศรษฐกิจและสังคม

2.3 งานวิจัยที่เกี่ยวข้อง

วิธีการวิเคราะห์ค่าเฉลี่ย ถูกนำมาประยุกต์ใช้ในงานวิจัยเพื่อวิเคราะห์และวัดผลการดำเนินงานของสถาบันการเงินอย่างธนาคารและสหกรณ์เป็นจำนวนมากเช่น Kabir และ Dey [9] ใช้วิธีการวิเคราะห์ค่าเฉลี่ยมาประยุกต์ใช้กับสถาบันการเงินอย่างธนาคารเพื่อเปรียบเทียบสถานการณ์ทางการเงินของธนาคาร 2 แห่งนั่นคือธนาคาร IFIC และ ธนาคาร EXIM โดยต้องการทราบว่าธนาคารทั้ง 2 แห่งนี้มีประสิทธิภาพทางการเงินเป็นอย่างไรตามหลักของหลักการวิเคราะห์ค่าเฉลี่ย Prasad และ Ravinder [10] นำเอาวิธีการวิเคราะห์ค่าเฉลี่ยมาวิเคราะห์และวัดผลการดำเนินงานของธนาคารในอินเดีย 20 แห่งและทำการเรียงลำดับธนาคารที่มีผลการดำเนินงานดีที่สุดจากมากไปน้อย วิการ์ตัน [11] ทำการประเมินผลการดำเนินงานของสหกรณ์ออมทรัพย์ครุระนอง จำกัด ในช่วงปี พ.ศ. 2549 - 2553 โดยการประยุกต์ใช้วิธีการวิเคราะห์ค่าเฉลี่ยร่วมกับข้อมูลและอัตราส่วนสำคัญของสหกรณ์เฉลี่ย ผลที่ได้คือช่วยให้ผู้ที่เกี่ยวข้องกับสหกรณ์เห็นถึงแนวโน้มผลดำเนินงานในแต่ละปีของสหกรณ์ และสามารถนำเอาข้อมูลที่ได้จากการวิเคราะห์ไปกำหนดทิศทางของสหกรณ์ต่อไป

3. วิธีการดำเนินการวิจัย

ในการวิเคราะห์ภาวะทางการเงินของสหกรณ์ โดยประยุกต์ใช้วิธีการวิเคราะห์ค่าเฉลี่ยและแสดงผลในรูปแบบธุรกิจอัจฉริยะ แบ่งออกเป็น 3 ขั้นตอนคือ การเตรียมข้อมูล การสร้างแบบจำลองการวิเคราะห์ภาวะทางการเงินโดยการประยุกต์ใช้วิธีการวิเคราะห์ค่าเฉลี่ยและการวิเคราะห์ภาวะทางการเงินของสหกรณ์ด้วยวิธีการวิเคราะห์ค่าเฉลี่ย โดยมีขั้นตอนการเตรียมข้อมูลและการสร้างแบบจำลอง ดังภาพที่ 2



ภาพที่ 2 : การสร้างแบบจำลองการวิเคราะห์ภาวะทางการเงิน

3.1 การเตรียมข้อมูล (Data Preparation)

จากภาพที่ 2 คณะผู้วิจัยได้รวบรวมข้อมูลงบการเงินของสหกรณ์กรณีศึกษากรมตรวจบัญชีสหกรณ์ที่ผู้สอบบัญชีได้ทำการตรวจสอบงบการเงินแล้ว ตั้งแต่ปี พ.ศ. 2552 - 2556 และข้อมูลรายงานผลการวิเคราะห์งบการเงินของผู้สอบบัญชีปี พ.ศ. 2554 - 2555 โดยนำข้อมูลมาผ่านกระบวนการทำความสะอาดข้อมูล (Data Cleaning) เพื่อจัดข้อมูลที่ไมเกี่ยวข้อง รวมข้อมูลที่มีอยู่หลายปีให้เป็นข้อมูลชุดเดียวกัน (Data Integration) และปรับเปลี่ยนรูปแบบข้อมูลเพื่อจัดรูปแบบข้อมูลให้เหมาะสมต่อการนำไปใช้งาน (Data Transformation)

3.2 การสร้างแบบจำลองการวิเคราะห์ภาวะทางการเงินโดยการประยุกต์ใช้วิธีการวิเคราะห์ความผล

นำข้อมูลผลการวิเคราะห์งบการเงินจากรายงานผลการสอบบัญชีของผู้สอบบัญชีมาหากฎความสัมพันธ์ เพื่อหาอัตราส่วนสำคัญทางการเงินที่สามารถบอกถึงสภาพทางการเงินของแต่ละมิติของวิธีการวิเคราะห์ความผลได้ ผ่านโปรแกรมประยุกต์ Weka โดยเลือกใช้วิธีการเอพริออริ (Apriori) จากการหากฎความสัมพันธ์พบว่า อัตราส่วนสำรองต่อสินทรัพย์ อัตราค่าใช้จ่ายดำเนินงานต่อกำไร และอัตรารูทหนี้ที่ชำระได้ตามกำหนด เป็นอัตราส่วนสำคัญทางการเงินที่สามารถบอกสภาพทางการเงินของวิธีการวิเคราะห์ความผลเบื้องต้นได้ โดยตัวอย่างกฎความสัมพันธ์ที่ได้ เช่น ผลการวิเคราะห์อัตราส่วนสำรองต่อสินทรัพย์อยู่ในเกณฑ์ดี ผลการวิเคราะห์ภาวะทางการเงินในมิติที่ 1 ก็อยู่ในเกณฑ์ดีด้วย เป็นต้น โดยมีค่าสนับสนุนเท่ากับ 0.1 และค่าความเชื่อมั่นเท่ากับ 0.95 ซึ่งมีสูตรการคำนวณดังสมการที่ 1 และ 2

$$\text{Support (XUY)} = P(XUY) \tag{1}$$

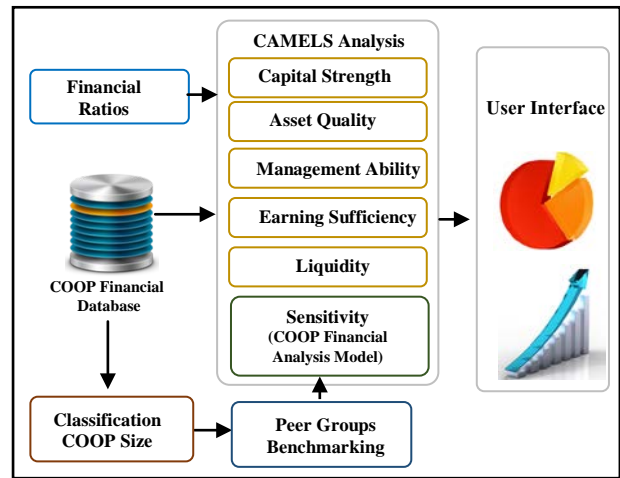
Support คือ ค่าความน่าจะเป็นของจำนวนของค่า X เกิดขึ้นคู่กับค่า Y ซึ่งมีค่าอยู่ระหว่าง [0-1] ถ้าเข้าใกล้ 1 หมายถึงเกิดรูปแบบนี้เป็นจำนวนมาก

$$\text{Confidence (X} \rightarrow \text{Y)} = \frac{P(XUY)}{P(X)} \tag{2}$$

Confidence คือ ค่าความน่าจะเป็นของจำนวนของค่า X เกิดขึ้นคู่กับค่า Y โดยที่มีลำดับเหตุการณ์เข้ามาเกี่ยวข้อง

จากนั้นนำข้อมูลผลการวิเคราะห์อัตราส่วนทางการเงินจากรายงานผลการสอบบัญชี โดยเลือกเฉพาะข้อมูลผลการวิเคราะห์ของอัตราส่วนสำรองต่อสินทรัพย์ อัตราค่าใช้จ่ายดำเนินงานต่อกำไรใน และอัตรารูทหนี้ที่ชำระได้ตามกำหนด มาทำการสร้างแบบจำลองการวิเคราะห์ภาวะทางการเงินโดยใช้การแบ่งประเภทข้อมูล ซึ่งคณะผู้วิจัยเลือกใช้วิธีการแบ่งประเภทข้อมูลแบบต้นไม้ตัดสินใจและใช้วิธีการ J48 จากนั้นตรวจสอบความถูกต้องของแบบจำลองด้วยแบบไขว้ 10 ส่วน (10-Folds Cross Validation) เพื่อหาระดับภาวะทางการเงินของสหกรณ์ และนำแบบจำลองที่ได้ไปใช้วิเคราะห์ในมิติที่ 6 ของการวิเคราะห์ความผล

3.3 การวิเคราะห์ภาวะทางการเงินของสหกรณ์ด้วยวิธีการวิเคราะห์ความผลและแบบจำลองการวิเคราะห์ภาวะทางการเงิน



ภาพที่ 3 : ขั้นตอนการวิเคราะห์ภาวะทางการเงินด้วยวิธีการวิเคราะห์ความผลและแบบจำลองการวิเคราะห์ภาวะทางการเงิน

จากภาพที่ 4 ทำการวิเคราะห์ภาวะทางการเงินของสหกรณ์จากงบการเงินของสหกรณ์ โดยนำข้อมูลงบการเงินของสหกรณ์ 5 ปีย้อนหลังมาคำนวณอัตราส่วนทางการเงินที่คณะผู้วิจัยได้ทำการศึกษาจากผู้เชี่ยวชาญแล้วพบว่าเหมาะที่จะนำมาใช้เป็นองค์ประกอบในการวิเคราะห์และวัดผลทางการเงินของสหกรณ์ในแต่ละมิติของวิธีการวิเคราะห์ความผล เพื่อให้สหกรณ์เห็นถึงภาวะทางการเงินของสหกรณ์ย้อนหลัง 5 ปี ว่ามีแนวโน้มเป็นอย่างไร ช่วยให้ผู้บริหารสหกรณ์สามารถวางแผนในการบริหารสหกรณ์ได้อย่างมีประสิทธิภาพ โดยมีรายละเอียดแต่ละมิติดังนี้

มิติที่ 1 ความเพียงพอของเงินทุนต่อความเสี่ยง

มิตินี้เป็นการวิเคราะห์ว่าสหกรณ์มีเงินทุนเพียงพอที่จะรองรับความเสี่ยงหรือไม่ และอัตราส่วนทางการเงินที่นำมาใช้ในการวิเคราะห์ในมิตินี้คือ อัตราส่วนหนี้สินต่อทุนสำรองต่อสินทรัพย์ การเติบโตของทุนสหกรณ์ การเติบโตของหนี้ และอัตราผลตอบแทนต่อส่วนของทุน ตัวอย่างการคำนวณอัตราส่วนทางการเงิน ดังสมการที่ 3

$$\text{Growth of Cooperative Capital} = \frac{\text{Capital Reserve}}{\text{Total Asset}} \tag{3}$$

มติที่ 2 คุณภาพของสินทรัพย์

คุณภาพของสินทรัพย์เป็นการวิเคราะห์ว่าสินทรัพย์ที่ลงทุนได้ก่อให้เกิดรายได้แก่สหกรณ์อย่างไร อัตราส่วนทางการเงินที่นำมาใช้ในการวิเคราะห์คือ อัตราค้างชำระของลูกหนี้ อัตราหมุนของสินทรัพย์ ผลตอบแทนต่อสินทรัพย์ และอัตราการเติบโตของสินทรัพย์ ตัวอย่างการคำนวณอัตราส่วนทางการเงินดังสมการที่ 4

$$\text{Accrued Liabilities} = \frac{\text{Debt on Schedule} \times 100}{\text{Maturity Debt}} \quad (4)$$

มติที่ 3 วัดความสามารถในการบริหาร

เป็นการวิเคราะห์ถึงความสามารถของฝ่ายบริหารในการวางกลยุทธ์ และการจัดการ โครงสร้างของสหกรณ์ โดยอัตราส่วนทางการเงินที่ใช้ในการวิเคราะห์คือ อัตราการเติบโตของธุรกิจ

มติที่ 4 การทำกำไร

มตินี้เป็น การวิเคราะห์ถึงคุณภาพและแนวโน้มการทำกำไรของสหกรณ์ใช้อัตราส่วนทางการเงินในการวิเคราะห์คืออัตรากำไรต่อสมาชิก เงินออมต่อสมาชิก หนี้สินต่อสมาชิก ค่าใช้จ่ายดำเนินงานต่อกำไรก่อนหัก และการเติบโตกำไร ตัวอย่างการคำนวณอัตราส่วนทางการเงิน ดังสมการที่ 5

$$\text{Expenses to Profit before} = \frac{\text{Expenses}}{\text{Profit Before Expense}} \quad (5)$$

มติที่ 5 สภาพคล่อง

สภาพคล่องเป็นการวัดความเพียงพอต่อความต้องการใช้เงินของสหกรณ์ โดยวัดได้จาก อัตราหมุนเวียน อัตราหมุนของสินค้า อายุเฉลี่ยของสินค้า และอัตราลูกหนี้ระยะสั้นที่ชำระ ตัวอย่างการคำนวณอัตราส่วนทางการเงิน ดังสมการที่ 6

$$\frac{\text{Short Term Debtor Paid on Schedule}}{\text{Short Term Maturity Debtor}} = \frac{\text{Short Term Debtor Paid on Schedule} \times 100}{\text{Short Term Maturity Debtor}} \quad (6)$$

มติที่ 6 ผลกระทบต่อธุรกิจ

เป็นการวิเคราะห์ปัจจัยเสี่ยงที่มีผลต่อการดำเนินงานของสหกรณ์ ในมตินี้คณะผู้วิจัยนำเอาแบบจำลองการวิเคราะห์ภาวะทางการเงินที่สร้างขึ้นมาทำการวิเคราะห์ในมตินี้เพื่อสรุปภาพรวมของภาวะทางการเงินของสหกรณ์ว่ามีความเสี่ยงที่ต้องเฝ้าระวังมากน้อยเพียงใด โดยมีระดับการเฝ้าระวัง 4 ระดับ คือ ระดับปกติ เฝ้าระวังมากขึ้น เฝ้าระวังพิเศษ และระวังเร่งด่วนพิเศษ

จากนั้นคณะผู้วิจัยดำเนินการสร้างเกณฑ์ในการเทียบเคียงทางการเงินภายในกลุ่มประเภทและขนาดสหกรณ์เดียวกัน เพื่อให้สหกรณ์สามารถเทียบเคียงผลการวิเคราะห์ภาวะทางการเงินกับสหกรณ์อื่น ๆ ในกลุ่มเดียวกันได้ โดยทำการจัดกลุ่มขนาดของสหกรณ์ จากข้อมูลทุนดำเนินงาน รายได้ และจำนวนสมาชิก ด้วยวิธีการแบ่งประเภทข้อมูล ซึ่งสามารถแบ่งออกได้ 4 ขนาดคือ เล็ก กลาง ใหญ่ และใหญ่มาก จากนั้นคำนวณค่าข้อมูลและอัตราส่วนสำคัญของสหกรณ์เฉลี่ยจากผลการคำนวณอัตราส่วนทางการเงินของวิธีการวิเคราะห์ความเล็ง 5 มิติ โดยมีสูตรการคำนวณดังสมการที่ 7

$$\text{Peer Group} = \frac{\sum_{i=1}^n X_i}{N} \quad (7)$$

ทำการแสดงผลการวิเคราะห์ภาวะทางการเงินในรูปแบบของธุรกิจอัจฉริยะ เพื่อให้ผู้ใช้งานสามารถเข้าใจผลการวิเคราะห์ได้ง่าย สะดวกและนำผลการวิเคราะห์ไปช่วยในการตัดสินใจได้อย่างรวดเร็ว

4. ผลการดำเนินงาน

4.1 ผลการสร้างแบบจำลอง

ตรวจสอบความถูกต้องของแบบจำลองการวิเคราะห์ภาวะทางการเงิน ด้วยวิธีตรวจสอบความถูกต้องแบบไขว้ 10 ส่วน (10-Fold Cross Validation) พบว่าแบบจำลองให้ค่าความแม่นยำสูงถึง 98.71% ดังตารางที่ 1

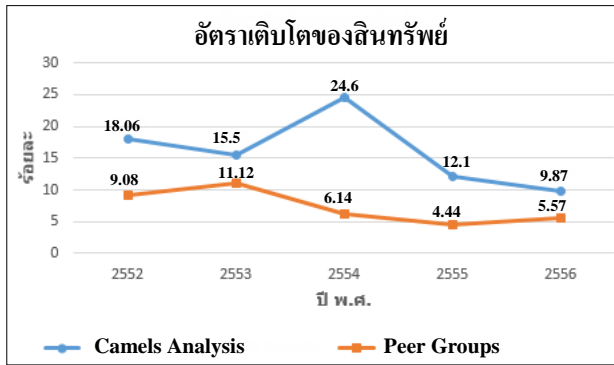
ตารางที่ 1: ผลการทดสอบความถูกต้องของแบบจำลอง

ผลการทดสอบ	จำนวนข้อมูล	ความถูกต้อง (%)
ความถูกต้องในการจัดระดับการเฝ้าระวัง	11,252	98.71
ความผิดพลาดในการจัดระดับการเฝ้าระวัง	146	1.28

4.2 ผลการพัฒนาการวิเคราะห์ภาวะทางการเงิน

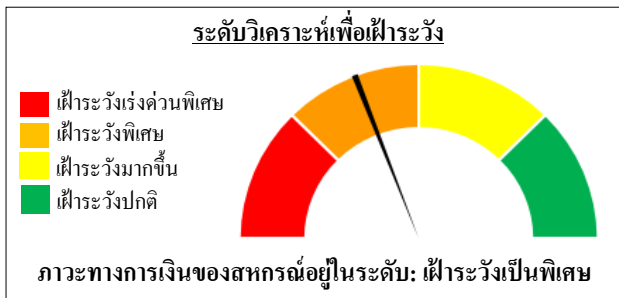
ผลการวิเคราะห์ภาวะทางการเงินของสหกรณ์โดยการประยุกต์ใช้วิธีการวิเคราะห์ความเล็ง สามารถแสดงผลการวิเคราะห์ได้ 4 ระดับ คือ ระดับวิเคราะห์เพื่อเฝ้าระวัง ระดับอัตราส่วนสำคัญทางการเงินของสหกรณ์ ระดับการวิเคราะห์ภาวะทางการเงินรายมิติของความเล็ง และระดับการวิเคราะห์เชิงลึกรายอัตราส่วนทางการเงินซึ่งเทียบเคียงกับข้อมูล 5 ปี ย้อนหลัง พร้อมให้ข้อเสนอแนะในการปรับปรุงพัฒนาการ

บริหารจัดการให้ประสบความสำเร็จ โดยแสดงผลในรูปแบบของกราฟ ดังภาพที่ 4



ภาพที่ 4: หน้าจอแสดงผลการวิเคราะห์ระดับรายอัตราส่วนทางการเงิน

หน้าจอแสดงผลการวิเคราะห์ภาวะทางการเงินระดับวิเคราะห์เพื่อเฝ้าระวังแสดงได้ดังภาพที่ 5



ภาพที่ 5: หน้าจอแสดงผลการวิเคราะห์ระดับวิเคราะห์เพื่อเฝ้าระวัง

5. สรุปผลการวิจัย และข้อเสนอแนะ

งานวิจัยนี้เสนอการวิเคราะห์ภาวะทางการเงินของสหกรณ์ โดยประยุกต์ใช้วิธีการวิเคราะห์ค่าเมตริกซึ่งเป็นตัววัดผลการดำเนินงานที่มีมาตรฐานและเชื่อถือได้ ร่วมกับการหาความสัมพันธ์และการแบ่งประเภทข้อมูลเพื่อสร้างแบบจำลองในการวิเคราะห์ภาวะทางการเงินของสหกรณ์ ผลการวิจัยพบว่าแบบจำลองที่สร้างขึ้นสามารถวิเคราะห์ภาวะทางการเงินของสหกรณ์ได้อย่างมีประสิทธิภาพและรวดเร็ว โดยมีค่าความแม่นยำของแบบจำลองจากการทดสอบแบบความถูกต้องแบบไขว้ 10 ส่วนเท่ากับ 98.71% ผู้บริหารสหกรณ์สามารถตรวจสอบสภาพทางการเงินและวิเคราะห์ระดับภาวะทางการเงินเพื่อเฝ้าระวังทางการเงินได้จริง

งานวิจัยในอนาคตสามารถนำเอาเทคนิคที่ช่วยในการพยากรณ์อย่าง Neural Network หรือ Naïve Bayes มาใช้ร่วมกับแบบจำลองการวิเคราะห์ภาวะทางการเงินของสหกรณ์ที่พัฒนาขึ้น เพื่อเพิ่มประสิทธิภาพในการวิเคราะห์แบบจำลอง

ทำให้ผู้บริหารสหกรณ์รับรู้สถานการณ์ทางการเงินได้อย่างต่อเนื่องทั้งในปัจจุบันและอนาคต ช่วยให้สามารถวางแผนสร้างความปลอดภัยหรือลดความเสี่ยงในการบริหารทางการเงินของสหกรณ์ได้ดียิ่งขึ้น

เอกสารอ้างอิง

- [1] สมเดช โรจน์บุรีเสถียร, *วิเคราะห์ทั้งการเงิน มองอดีต ปัจจุบัน อนาคต*, ธรรมนิติ เพลส: กรุงเทพมหานคร, 2552.
- [2] กรมตรวจบัญชีสหกรณ์, *คู่มือการประยุกต์ใช้ CAMELS ANALYSIS วิเคราะห์ภาวะเศรษฐกิจทางการเงินของสหกรณ์และกลุ่มเกษตรกร*, แอร์บอร์น พรินต์: กรุงเทพมหานคร, 2549.
- [3] A. Sarker, "CAMEL Rating System in the Context of Islamic Banking: A Proposed 'S' for Shariah Framework" *Journal of Islamic Economics and Finance*, vol. 1, no.1, pp. 78-84, 2005.
- [4] กรมตรวจบัญชีสหกรณ์, *ข้อมูลและอัตราส่วนสำคัญของสหกรณ์ และกลุ่มเกษตรกรเฉลี่ย*, แอร์บอร์น พรินต์: กรุงเทพมหานคร, 2552
- [5] M. Z. Elbashir, P. A. Collier, and M. J. Davern, "Measuring the effects of business intelligence systems: The relationship between business process and organizational performance," *International Journal of Accounting Information Systems*, vol. 9, pp. 135-153, 2008.
- [6] พัศกร สิงห์โต, อัครวุฒิ ประมะปัญญา, ปฐมาภรณ์ เถาว์พัน "การเปรียบเทียบประสิทธิภาพการจำแนกข้อมูลด้วยเทคนิคการทำเหมืองข้อมูล". *National Conference on Computer Information Technologies*, 2011, pp. 42-48.
- [7] D. M. Farid, et al, "Hybrid decision tree and naïve Bayes classifiers for multi-class classification tasks" *Expert Systems with Applications*. Vol 41, No. 4, pp. 1937-1946, 2014
- [8] ชีระวัฒน์ แสนปัญญา. "การพยากรณ์การซื้อสินค้าประเภทกระเบื้องปูพื้นสำหรับร้านค้าปลีกโดยวิธีกฎความสัมพันธ์". *การประชุมเสนอผลงานวิจัยระดับบัณฑิตศึกษา มสธ. ครั้งที่ 1*, 2554
- [9] A. Kabir, and S. Dey, "Performance Analysis through CAMEL Rating: A Comparative Study of Selected Private Commercial Banks in Bangladesh" *Journal of Politics & Governance*, Vol. 1, pp. 16-25, 2012.
- [10] K. V. N. Prasad, and G. A. Ravinder, "Camel Model Analysis of Nationalized Banks in India" *International Journal of Trade and Commerce-IIARTC*, Vol. 1, pp. 23-33, 2012.
- [11] วิภารัตน์ สอนวิสัย. "การประเมินผลการดำเนินงานของสหกรณ์ : กรณีศึกษา สหกรณ์ออมทรัพย์นครระนอง", *การค้นคว้าแบบอิสระ บริหารธุรกิจมหาบัณฑิต*, มหาวิทยาลัยเทคโนโลยี ราชมงคล, 2554

ระบบสมาชิกสำหรับแอนดรอยด์สมาร์ทโฟนโดยใช้เอ็นเอฟซีการ์ดอิมูเลชันโหมด

A Membership System for Android Smartphone with NFC Card Emulation Mode

ณัฐพล แซ่ลิม¹ (Nattaphon Saelim)¹ และเชษฐ พัฒโนทัย (Chate Patanothai)²

¹ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

Nattaphon.Sa@student.chula.ac.th, chate.p@chula.ac.th

บทคัดย่อ

ปัญหาความยุ่งยากและซับซ้อนที่เกิดจากการส่งข้อมูลแบบเพียร์ทูเพียร์บนแอนดรอยด์สมาร์ทโฟนซึ่งไม่อำนวยความสะดวกในการใช้งาน โครงการนี้จึงนำเสนอการปรับปรุงประสิทธิภาพการส่งข้อมูลด้วยวิธีเอ็นเอฟซีการ์ดอิมูเลชันในขั้นตอนการตรวจสอบสมาชิก เพื่อลดขั้นตอนการส่งข้อมูลที่ไม่จำเป็น ผลการทดลองด้วยวิธีใหม่นี้ระบบสามารถส่งข้อมูลได้เร็วขึ้นกว่า เดิมถึง 30.28% ต้นแบบของระบบดังกล่าวสามารถนำไปประยุกต์เข้ากับระบบสมาชิกต่างๆ รวมทั้งระบบที่ต้องการความรวดเร็วในการใช้งาน เช่นระบบโดยสารรถไฟฟ้า BTS, MRT ได้เป็นต้น

คำสำคัญ: เอ็นเอฟซี ระบบสมาชิก โฮสต์เบสการ์ดอิมูเลชัน

Abstract

Sending information via peer-to-peer on Android Smartphones is complex and inconvenient. This project proposed a way to verify membership information via NFC card emulation as an alternative way to peer-to-peer method. By using this method, the efficiency has improved by 30.28% of the original speed, since the unnecessary steps are reduced. This implementation of NFC emulation can be applied to many membership systems and other system especially those require rapid access time such as BTS and MRT ticket systems.

Keywords: NFC, Membership System, Host-based Card Emulation

1. บทนำ

ในการแข่งขันทางธุรกิจ ร้านค้าต่าง ๆ ให้ความสำคัญกับการพัฒนาด้านการจัดการลูกค้าสัมพันธ์ โดยมีวัตถุประสงค์หลักคือ นำเสนอสินค้าและบริการที่สร้างความสุข เพื่อให้ลูกค้าใช้สินค้าเหล่านั้นอย่างสม่ำเสมอ กลยุทธ์ต่าง ๆ ที่ใช้เพื่อสร้างสัมพันธ์ที่ดีกับลูกค้า หนึ่งในนั้นคือ การทำบัตรสมาชิกเพิ่มสิทธิประโยชน์หรือส่วนลดให้กับลูกค้า ด้วยเหตุนี้เองร้านค้าต่าง ๆ จึงออกบัตรสมาชิกให้กับลูกค้าเป็นจำนวนมาก ส่งผลให้ลูกค้าต้องพกบัตรสมาชิกของร้านค้าต่าง ๆ ซึ่งไม่อำนวยความสะดวกในการพกพา ผู้ทำโครงการนี้ได้พัฒนาต้นแบบ (Prototype) ของระบบสมาชิกซึ่งใช้เทคโนโลยีเอ็นเอฟซี (Near Field Communication: NFC) ที่ผนวกเข้ากับแอนดรอยด์สมาร์ทโฟน ลูกค้าสามารถใช้สมาร์ทโฟนทดแทนการพกพาบัตรสมาชิกของร้านค้าต่าง ๆ ได้ อย่างไรก็ตามระบบต้นแบบแรกที่ได้พัฒนาขึ้นนั้นยังพบปัญหาอยู่มาก เนื่องจากระบบเดิมที่พัฒนาขึ้นใช้รูปแบบการส่งข้อมูลแบบเพียร์ทูเพียร์ ซึ่งการทำงานดังกล่าวมีขั้นตอนที่ยุ่งยากและซับซ้อน ไม่อำนวยความสะดวกในการใช้งาน

โครงการนี้จึงนำเสนอต้นแบบของระบบสมาชิกโดยมุ่งเน้นปรับปรุงประสิทธิภาพการทำงานของระบบเดิม ในขั้นตอนการตรวจสอบข้อมูลสมาชิก โดยใช้รูปแบบการส่งข้อมูลแบบเอ็นเอฟซีการ์ดอิมูเลชัน ซึ่งเป็นรูปแบบการส่งข้อมูลที่เพิ่มเข้ามาใหม่ในระบบปฏิบัติการแอนดรอยด์เวอร์ชัน 4.4 ขึ้นไป [10] มีความแตกต่างจากการส่งข้อมูลแบบเพียร์ทูเพียร์แบบเดิมสามารถทำหน้าที่เปรียบเสมือนบัตรสมาชิกของร้านค้าต่าง ๆ ได้ดียิ่งขึ้น ลดขั้นตอนที่ซับซ้อนและเพิ่มความสะดวกในการใช้งาน

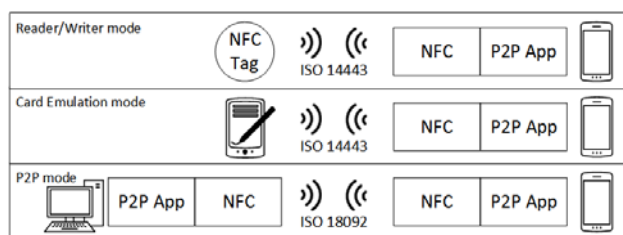
2. ทฤษฎีที่เกี่ยวข้อง

2.1 แอนดรอยด์

แอนดรอยด์ [6] เป็นระบบปฏิบัติการโดยมีพื้นฐานอยู่บนระบบปฏิบัติการลินุกซ์ ถูกออกแบบสำหรับอุปกรณ์พกพาขนาดเล็กที่ใช้จอสัมผัสโดยเฉพาะ เช่น สมาร์ทโฟน และแท็บเล็ต ริเริ่มคิดค้นและพัฒนาจากบริษัทแอนดรอยด์ (Android, Inc.)

2.2 เทคโนโลยีเอ็นเอฟซี

เทคโนโลยีเอ็นเอฟซีเป็นเทคโนโลยีสื่อสารไร้สายระยะใกล้สามารถรองรับการสื่อสารสองทางระหว่างเครื่องมืออิเล็กทรอนิกส์ในระยะใกล้ประมาณ 1 - 4 ซม. ทำงานได้ดีกับโครงสร้างพื้นฐานแบบไร้สัมผัสบนคลื่นความถี่ 13.56 MHz. รับส่งข้อมูลด้วยความเร็ว 424 Kbps บนพื้นฐานมาตรฐานไอเอสโอ/ไออีซี 18092 NFC IP-1 [4] และไอเอสโอ/ไออีซี 14443 [5] (Philips MIFARE and Sony's FeliCa) โดยมาตรฐานดังกล่าวได้เสนอโหมดการทำงานทั้งสามแบบ [2] ที่แตกต่างกันดังภาพที่ 1



ภาพที่ 1: โหมดทำงานของเอ็นเอฟซีและมาตรฐาน

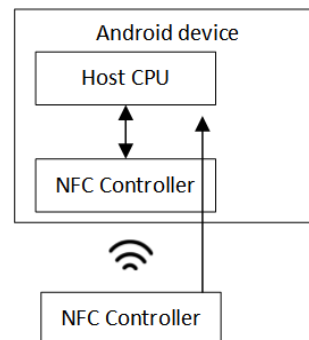
2.2.1 อ่าน/เขียน (Reader/Writer mode)

โหมดนี้ อุปกรณ์เอ็นเอฟซีสามารถทำตัวเสมือนเป็นเครื่องอ่าน และเขียนบัตรสมาร์ตการ์ดแบบไร้สัมผัส (Contactless Smart Card) หรือเรียกอีกอย่างว่าแท็ก (Tag) สามารถอ่านข้อมูลจากแท็กที่อยู่ในบัตร หรือติดอยู่ในจุดให้บริการข้อมูลได้ (Smartposter) โหมดดังกล่าวสอดคล้องกับมาตรฐานไอเอสโอ/ไออีซี 14443

2.2.2 เอ็นเอฟซีการ์ดอิมูเลชัน (NFC Card Emulation Mode)

โหมดนี้จะทำงานเสมือนเป็นบัตรสมาร์ตการ์ดแบบไร้สัมผัส [7] ซึ่งรับส่งข้อมูลบนพื้นฐานมาตรฐานไอเอสโอ/ไออีซี

14443-4 (ISO-DEP) และแอปพลิเคชันโพรโทคอลดาต้ายูนิท [8] (Application Protocol Data Units: APDUs) โพรโทคอลดังกล่าวสามารถทำหน้าที่สื่อสารระหว่างเครื่องอ่านบัตรสมาร์ตการ์ด และบัตรสมาร์ตการ์ด ภาพที่ 2 แสดงถึงการทำงานในโหมดการ์ดอิมูเลชัน



ภาพที่ 2: การทำงานในโหมดการ์ดอิมูเลชัน

ซึ่งอุปกรณ์ที่สอดคล้องตามมาตรฐานเอ็นเอฟซี สามารถทำตัวเป็นบัตรสมาร์ตการ์ดแบบไร้สัมผัส เพื่อใช้ในการทำธุรกรรมต่าง ๆ ได้ อย่างไรก็ตามระบบปฏิบัติการแอนดรอยด์ที่เวอร์ชันต่ำกว่า 4.4 ไม่สามารถใช้งานโหมดดังกล่าวได้ [10]

2.2.3 เพียร์ทูเพียร์ (Peer-to-Peer Mode)

โหมดนี้จะทำการแลกเปลี่ยนข้อมูลระหว่างอุปกรณ์เอ็นเอฟซีด้วยกันเช่น ข้อมูลนามบัตร รูปถ่าย หรือเพิ่มข้อมูลอื่น ๆ เป็นต้น ซึ่งโหมดดังกล่าวสอดคล้องกับมาตรฐานไอเอสโอ/ไออีซี 18092

3. วรรณกรรมที่เกี่ยวข้อง

โปรแกรมประยุกต์แสตมป์ (Stamp Application) [12] เป็นโปรแกรมที่ใช้สำหรับสะสมแสตมป์ เพื่อรับส่วนลดหรือสิทธิประโยชน์จากทางร้านค้าต่าง ๆ โดยโปรแกรมดังกล่าวทำงานร่วมกับฮาร์ดแวร์พิเศษ อย่างไรก็ตามฮาร์ดแวร์ดังกล่าวมีราคาค่อนข้างสูง โปรแกรมประยุกต์กอติต (Gotit Application) [13] เป็นโปรแกรมที่ใช้สำหรับสะสมแสตมป์ เพื่อรับส่วนลดหรือสิทธิประโยชน์จากทางร้านค้าต่าง ๆ ที่ทำงานร่วมกับรหัสคิวอาร์ (QR Code) อย่างไรก็ตามโปรแกรมดังกล่าวไม่สามารถอ่านรหัสคิวอาร์ในสถานะที่มีแสงน้อย หรือได้ไม่ดีเท่าที่ควร

ตารางที่ 1 แสดงการเปรียบเทียบคุณสมบัติของระบบสมาชิก และ โปรแกรมประยุกต์

ตารางที่ 1: เปรียบเทียบคุณสมบัติของระบบสมาชิก / โปรแกรม

คุณสมบัติของระบบ	สมาชิก	แสดมป์	กอดติด
ทำงานได้ในภาวะแสงน้อย	✓	✓	✗
ฮาร์ดแวร์ราคาถูก	✓	✗	✓

งานวิจัย [1] ได้เสนอระบบต้นแบบสำหรับการชอปปิ้งในห้างสรรพสินค้าด้วยเทคโนโลยีเอ็นเอฟซีบนแอนดรอยด์แพลตฟอร์ม ผู้วิจัยทำการส่งข้อมูลรายละเอียดสินค้าโดยใช้การส่งข้อมูลแบบเพียร์ทูเพียร์ ซึ่งมีข้อดีคือสามารถรับส่งข้อมูลรายละเอียดของสินค้าแม้ในสภาวะออฟไลน์ได้ งานวิจัย [2] ได้นำเสนอระบบการชำระเงินที่มีความปลอดภัยบนเทคโนโลยีเอ็นเอฟซีด้วยแอนดรอยด์แพลตฟอร์มโดยใช้การส่งข้อมูลแบบเพียร์ทูเพียร์เช่นกัน งานวิจัยนี้มีข้อดีคือการรักษาความปลอดภัยของข้อมูลการเงิน สร้างความมั่นใจให้กับผู้ใช้งาน ทว่าจากงานวิจัยทั้ง 2 ผู้วิจัยใช้การส่งข้อมูลแบบเพียร์ทูเพียร์ในการรับส่งข้อมูล มีขั้นตอนยุ่งยากและซับซ้อน ไม่สะดวกในการใช้งาน ซึ่งในขณะนั้นระบบปฏิบัติการที่ผู้วิจัยใช้พัฒนาอยู่เวอร์ชันต่ำกว่า 4.4 นั้นยังไม่รองรับการส่งข้อมูลด้วยรูปแบบเอ็นเอฟซีการ์ดอิมูเลชันใหม่ที่มีความสะดวกในการใช้งานมากกว่า ผู้ทำโครงการจึงสังเกตเห็นว่าในระบบปฏิบัติการเวอร์ชันใหม่นี้ น่าจะเป็นแนวคิดที่เหมาะสมกับการนำเข้ามาปรับปรุงขั้นตอนการส่งข้อมูลให้มีประสิทธิภาพที่ดียิ่งขึ้น

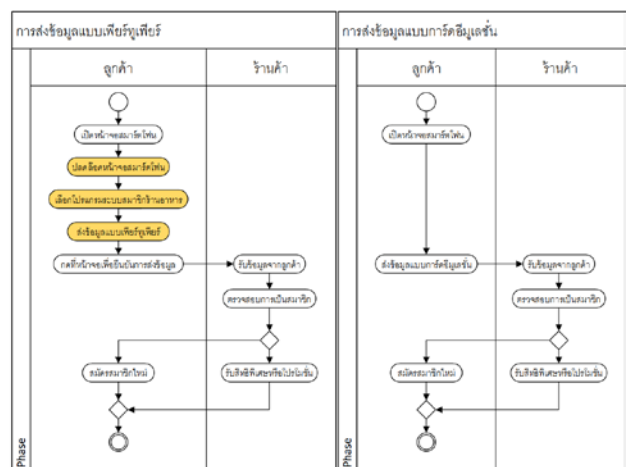
4. วิธีการดำเนินการวิจัย

4.1 ออกแบบแนวคิด

ผู้ทำโครงการคำนึงความสะดวกในการใช้งานสมาร์ทโฟนเพื่อทดแทนการบัตร เพื่อยืนยันว่าระบบที่พัฒนาขึ้นจะสามารถลดขั้นตอนที่ไม่จำเป็นสำหรับการส่งข้อมูลได้ จึงทำการปรับปรุงขั้นตอนการส่งข้อมูล พัฒนาโดยใช้วิธีโฮสต์เบสการ์ดอิมูเลชัน (Host-based Card Emulation: HCE) ซึ่งเป็นแอนดรอยด์เซอร์วิสที่ทำงานเป็นพื้นหลัง (Background Service) ขั้นตอนการสร้างโฮสต์เบสนั้นผู้ทำโครงการจะทำการสร้างคลาสที่ต่อส่วนขยายมาจากคลาส HostApuService และทำ

การอิมพลีเมนต์เมทอด processCommandApu และ onDeactivated ตามลำดับ จากนั้นเซอร์วิสที่สร้างขึ้นจะต้องถูกระบุไว้ในไฟล์แมนนิเฟส (Manifest file) เพื่อเป็นการประกาศว่าระบบดังกล่าวมีเซอร์วิสถูกติดตั้งอยู่รวมถึงต้องทำการเปิดสิทธิ์ของ “android.permission.BIND_NFC_SERVICE” เพื่อเป็นการอนุญาตให้เซอร์วิสที่สร้างขึ้นสามารถทำงานร่วมกับโปรแกรมภายนอกได้

จากนั้นทำการเปรียบเทียบขั้นตอนการตรวจสอบสมาชิกระหว่างการส่งข้อมูลแบบเพียร์ทูเพียร์ กับการส่งข้อมูลแบบการ์ดอิมูเลชัน ดังภาพที่ 3 จากภาพด้านซ้ายคือแผนภาพกิจกรรมการส่งข้อมูลแบบเพียร์ทูเพียร์ ด้านขวาคือการส่งข้อมูลด้วยวิธีการอิมูเลชัน จะเห็นได้ว่า การส่งข้อมูลแบบการ์ดอิมูเลชันนั้น ผู้ใช้งานไม่จำเป็นต้องปลดล็อคหน้าจอสมาาร์ทโฟน และไม่จำเป็นต้องค้นหาโปรแกรมที่ต้องการ รวมถึงกดหน้าจอเพื่อยืนยันการส่งข้อมูล เนื่องจากรูปแบบการส่งข้อมูลดังกล่าวมีการทำงานเป็นพื้นหลัง ส่งผลให้มีขั้นตอนการส่งข้อมูลที่น้อยกว่า (ดังที่แสดงในการเน้นสี)

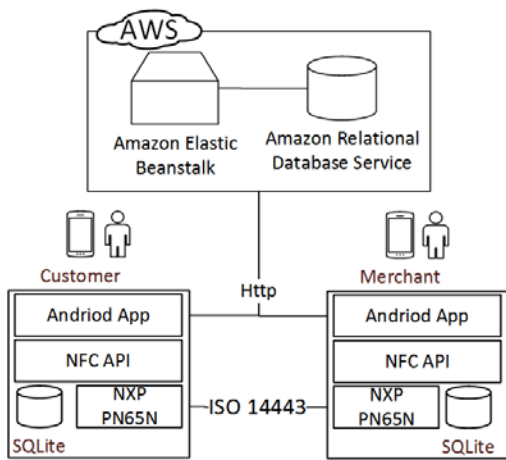


ภาพที่ 3: แผนภาพกิจกรรม เปรียบเทียบขั้นตอนการตรวจสอบสมาชิก โดยใช้การส่งข้อมูลแบบเพียร์ทูเพียร์ และแบบการ์ดอิมูเลชัน

4.2 ออกแบบสถาปัตยกรรมของระบบ

ต้นแบบของระบบสมาชิกที่ได้พัฒนาขึ้น ประกอบไปด้วยโปรแกรมประยุกต์แอนดรอยด์ 2 ส่วนทั้งในส่วนของลูกค้าและร้านค้า ผู้ใช้งานสามารถทำการสมัครสมาชิกได้โดยกรอกข้อมูลชื่อ นามสกุล อีเมล หมายเลขโทรศัพท์ และรหัสผ่าน โดย

โปรแกรมจะสร้างรหัสประจำตัวซอฟต์แวร์ขึ้น (Identifying App Installations) [11] เพื่อใช้ในการระบุตัวตนของลูกค้า จากนั้นข้อมูลทั้งหมดจะถูกส่งขึ้นไปบันทึกที่เซิร์ฟเวอร์ของ Amazon Elastic Beanstalk ซึ่งเป็นเวปเซอร์วิซกลุ่มเมฆ (Cloud Service) ที่พัฒนาขึ้นเพื่อรองรับการทำงานจากระบบสมาชิกร้านอาหาร โดยจะต่อกับส่วนของข้อมูล ซึ่งใช้ Amazon Relational Database Service ทำหน้าที่ในการเก็บข้อมูลต่าง ๆ เช่นข้อมูลสมาชิก ข้อมูลร้านค้า เป็นต้น ภาพที่ 4 แสดงถึงระบบสถาปัตยกรรมต้นแบบของระบบสมาชิก



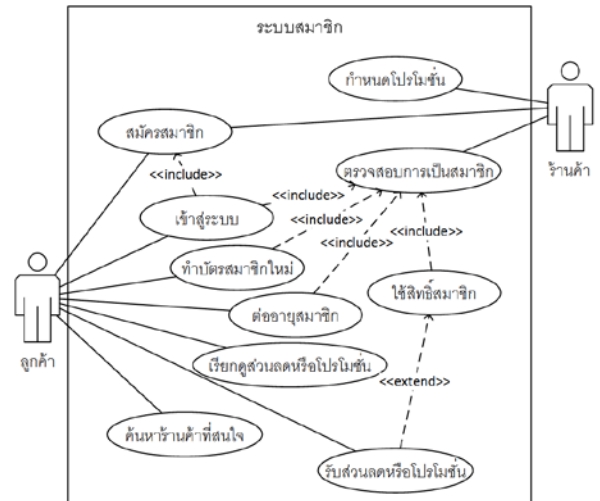
ภาพที่ 4: สถาปัตยกรรมของระบบสมาชิก

นอกจากนี้ระบบต้นแบบที่ได้พัฒนาขึ้นลูกค้าสามารถค้นหาข้อมูลร้านค้าที่สนใจ เรียกดูข้อมูลส่วนลด หรือโปรโมชั่นที่ต้องการได้ ลูกค้าสามารถทำการต่ออายุสมาชิก อีกทั้งในกรณีที่สมาร์ทโฟนเกิดสูญหาย ลูกค้าสามารถนำสมาร์ทโฟนเครื่องใหม่มาขอทำบัตรสมาชิกใหม่เพื่อเรียกข้อมูลลูกค้าเดิมกลับมาได้ดังภาพที่ 5 แสดงถึงแผนภาพยูสเคสของระบบสมาชิก และแผนภาพแสดงความสัมพันธ์ระหว่างข้อมูล (Entity Relationship Diagram) ดังภาพที่ 6

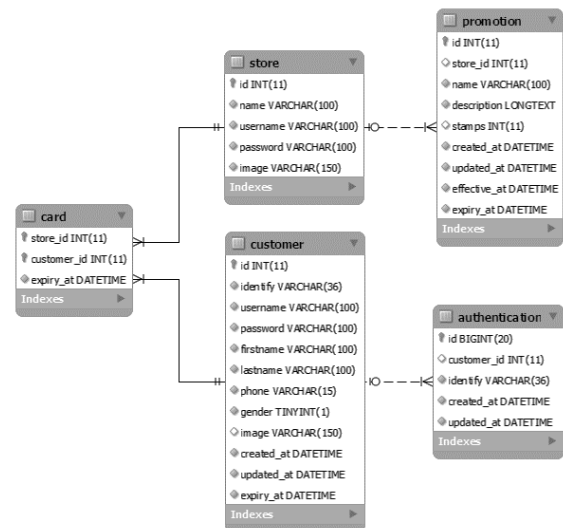
4.3 ออกแบบการทดลอง

ระบบต้นแบบที่ได้พัฒนาขึ้นถูกติดตั้งลงบนระบบปฏิบัติการแอนดรอยด์เวอร์ชัน 4.4 ขึ้นไปเพื่อรองรับเทคโนโลยีเอ็นเอฟซี และติดตั้งระบบต้นแบบฝั่งเซิร์ฟเวอร์ที่พัฒนาด้วยเฟรมเวิร์กซิมโฟนี (Symfony2) ภาษาพีเอชพี ในส่วนแรกผู้ทำโครงการจะใช้การทดสอบด้วยแบบจำลองคีส์โทรคเลเวลโมเดล

(Keystroke – Level Model: KLM) [9] เพื่อประมาณค่าเวลาและประสิทธิภาพในการใช้งานของระบบในขั้นตอนการส่งข้อมูล โดยแบบจำลองดังกล่าวสามารถประมาณค่าเวลาที่ใช้ในการทำกิจกรรมต่าง ๆ ได้ และใน ส่วนที่ 2 ได้จำลองสถานการณ์เมื่อลูกค้าเข้ามาใช้บริการร้านอาหารเป็นจำนวนมาก โดยจำลองปริมาณลูกค้าที่เข้ามาต่อคิวเพื่อรับบริการมีจำนวน 100 ราย และทำการเปรียบเทียบการส่งข้อมูลทั้ง 2 วิธี



ภาพที่ 5: แผนภาพยูสเคสของระบบสมาชิก



ภาพที่ 6: แผนภาพแสดงความสัมพันธ์ระหว่างข้อมูลของระบบสมาชิก

5. ผลการดำเนินงาน

5.1 ผลการทดลองการประมาณค่าเวลาด้วยแบบจำลองคีย์ส

โทรศ

ในการทดสอบนี้ ค่าที่นำมาใช้ในการประมาณคือเวลาการของการกดบนหน้าจอ “Finger Movement = f” เท่ากับ 0.23 วินาที เวลาของการเลื่อนเพื่อปลดล็อกหน้าจอหรือการแตะด้านหลังของเครื่องสมาร์ทโฟนกับอีกเครื่องเข้าด้วยกัน “Gestures = g” เท่ากับ 0.80 วินาที ตารางที่ 2 แสดงถึงเวลาที่ใช้ในขั้นตอนการตรวจสอบสมาชิกด้วยการส่งข้อมูลแบบวิธีต่างๆ จะเห็นได้ว่าวิธีการส่งข้อมูลแบบเพียร์ทูเพียร์ใช้เวลาานานกว่า เนื่องจากมีขั้นตอนที่ซับซ้อนและยุ่งยากกว่า

ตารางที่ 2: เปรียบเทียบเวลาที่ใช้ในการส่งข้อมูลทั้ง 2 รูปแบบ

ขั้นตอน	รูปแบบการส่ง / เวลาที่ใช้ (s)	
	เพียร์ทูเพียร์	การ์ดอีมีเลชั่น
เปิดหน้าจอสมาร์ทโฟน	0.80	0.80
ปลดล็อกหน้าจอสมาร์ทโฟน	0.80	-
เลือกโปรแกรมระบบสมาชิก (g ≤ 3)*	0.23	-
	+ (0.80 * 3)	
แตะด้านหลังของเครื่องสมาร์ทโฟนกับอีกเครื่องเข้าด้วยกัน	0.80	0.80
กดที่หน้าจอเพื่อยืนยันการส่งข้อมูล	0.23	-
รวมเวลา	5.26	1.60

* การเลื่อนหน้าจอจำนวน 3 ครั้งเป็นจำนวนที่เพียงพอต่อการค้นหาโปรแกรมที่สนใจ

5.2 ผลการทดลองด้านเวลาในการส่งข้อมูล

ผลการทดสอบเวลาที่ใช้วัดเป็นหน่วยมิลลิวินาที (Millisecond – ms) จำนวน 100 ครั้ง ได้ค่าเฉลี่ย (Average) ค่ามากที่สุด (Max) ค่าน้อยสุด (Min) ค่าเบี่ยงเบนมาตรฐาน (Standard Deviation - SD) ดังแสดงตารางที่ 3

เพื่อหาคำตอบของสมมุติฐานว่าความแตกต่างของรูปแบบการส่งข้อมูลมีผลต่อความเร็วที่ได้หรือไม่ ผู้ทำโครงการทำการ

ตั้งสมมุติฐานหลัก (null hypothesis – H₀) และสมมุติฐานรอง (alternative hypothesis – H₁) ไว้ดังนี้

H₀ : ความเร็วในการส่งข้อมูลไม่ต่างแตกต่างกันเมื่อรูปแบบการส่งข้อมูลต่างกัน

H₁ : ความเร็วในการส่งข้อมูลต่างแตกต่างกันเมื่อรูปแบบการส่งข้อมูลต่างกัน

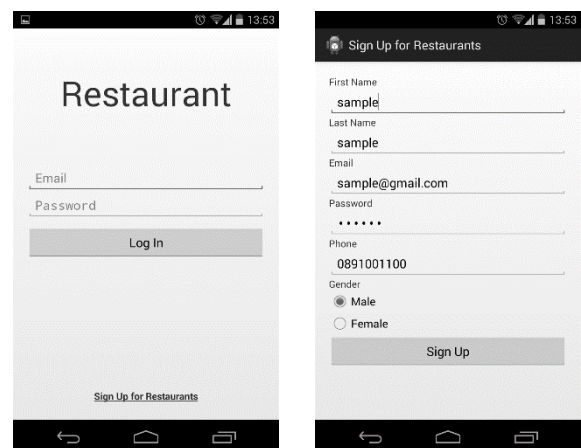
ใช้การทดสอบของ Student's t-test หาค่าความน่าจะเป็นแบบแจกแจง 2 ด้าน โดยกำหนดระดับนัยสำคัญ (level of significant – α) ไว้ที่ 0.1 ค่า P-value มีค่าเท่ากับ 6.52348E-12 ซึ่งต่ำกว่าค่าระดับนัยสำคัญ จึงสรุปผลว่า ปฏิเสธ H₀ ซึ่งก็คือความเร็วในการส่งข้อมูลแตกต่างกันเมื่อรูปแบบการส่งต่างกัน

ตารางที่ 3: ผลทดสอบเวลาที่ใช้นรูปแบบการส่งข้อมูลทั้ง 2 รูปแบบ

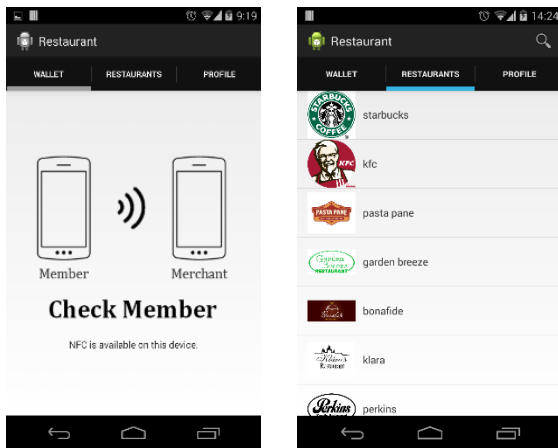
รูปแบบการส่งข้อมูล	เพียร์ทูเพียร์	การ์ดอีมีเลชั่น	
เวลา (ms)	Average	1816.86	2606.25
	Max	5392	8782
	Min	1351	1863
	SD	478.32	887.14

5.3 ตัวอย่างหน้าจอระบบต้นแบบ

หน้าจอหลักของโปรแกรมประยุกต์แอนดรอยด์ ประกอบไปด้วยหน้าจอการการเข้าสู่ระบบ หน้าจอการสมัครสมาชิกของลูกค้า หน้าจอรอเพื่อส่ง – รับข้อมูล และหน้าจอสำหรับเรียกดูข้อมูลร้านค้า แสดงดังภาพที่ 7 และภาพที่ 8 ตามลำดับ



ภาพที่ 7: หน้าจอการเข้าสู่ระบบและการสมัครสมาชิก



ภาพที่ 8: หน้าจอการรอเพื่อรับ – ส่งข้อมูลและการเรียกดูข้อมูลร้านค้า

6. สรุป

การขั้นตอนการตรวจสอบสมาชิกโดยใช้การส่งข้อมูลแบบการ์ดอิมูเลชันนั้น มีขั้นตอนการทำงานที่น้อยกว่า และใช้ระยะเวลาที่สั้นกว่า ประโยชน์ที่ได้สามารถนำไปปรับปรุงประสิทธิภาพของระบบสมาชิก ซึ่งจะช่วยอำนวยความสะดวกในการใช้สมาร์ทโฟนแทนการพกบัตรสมาชิกได้ดียิ่งขึ้น ในการพัฒนาครั้งต่อไป ระบบดังกล่าวสามารถนำไปประยุกต์เข้ากับระบบต่าง ๆ ที่ต้องการความรวดเร็ว เช่นระบบโดยสารรถไฟฟ้า BTS, MRT สามารถใช้แทนบัตรรถไฟฟ้าในการเข้าออกสถานีได้ เป็นต้น

7. กิตติกรรมประกาศ

ขอพระคุณ อ.เชษฐ พัฒโนทัย อาจารย์ที่ปรึกษาโครงการ และ ผศ.ดร.เกริก ภิรมย์โสภานเป็นอย่างยิ่งที่ได้สละเวลาให้คำปรึกษา คำแนะนำ และแนวทางสำหรับการทำโครงการ รวมทั้งเป็นผู้ประสานงานให้ความช่วยเหลือต่าง ๆ

เอกสารอ้างอิง

- [1] Husni, E.; Purwanto, S., "Shopping application system with Near Field Communication (NFC) based on Android," *System Engineering and Technology (ICSET), 2012 International Conference on*, vol., no., pp.1,6, 11-12 Sept. 2012.
- [2] Mainetti, L.; Patrono, L.; Vergallo, R., "IDA-Pay: An innovative micro-payment system based on NFC technology for Android mobile devices," *Software,*

Telecommunications and Computer Networks (SoftCOM), 2012 20th International Conference on, vol., no., pp.1,6, 11-13 Sept. 2012.

- [3] Divyan M. Konidala, et al., "Resuscitating Privacy-Preserving Mobile Payment with Customer in Complete Control," *Journal of Personal and Ubiquitous Computing (PUC)*. Vol.16, pp.643-654, 2012.
- [4] L. Catarinucci, S. Tedesco, D. De Donno, L. Tarricone: "Platform-Robust Passive UHF RFID Tags: a Case-Study in Robotics," *Progress In Electromagnetics Research C*, Vol. 30, 27-39, 2012.
- [5] International Standard ISO/IEC 14443-1-2-3-4, Identification cards - Contactless integrated circuit cards - Proximity cards, 2008-07-15, ISO/IEC 2008, Switzerland.
- [6] Android operating system, (2013, September 6). Available: [http://en.wikipedia.org/wiki/Android_\(operating_system\)](http://en.wikipedia.org/wiki/Android_(operating_system))
- [7] Android KitKat, (2014, January 30) Available: <http://developer.android.com/about/versions/kitkat.html>
- [8] Smart card application protocol data unit, (2014, January 30) Available: http://en.wikipedia.org/wiki/Smart_card_application_protocol_data_unit
- [9] Holleis, P. et al. Keystroke-Level Model for Advanced Mobile Phone Interaction. To appear in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems 2007 (CHI '07)*, San Jose, CA, USA, April/May 2007.
- [10] Host-based Card Emulation, (2014, Mar 10) Available: <http://developer.android.com/guide/topics/connectivity/nfc/hce.html>
- [11] Identifying App Installations, (2014, Mar 10) Available: <http://android-developers.blogspot.in/2011/03/identifying-app-installations.html>
- [12] Mobiliti Co., Ltd., (2013, September 6). *Stamp Application*. Available: <http://www.getmystamp.com/>
- [13] Got Apps Co., Ltd., (2013, September 6). *Gotit Application*. Available: <https://www.you-got.it/>

การศึกษาผลการเรียนรู้ของผู้เรียนที่ใช้กระบวนการเรียนรู้แบบโครงการร่วมกับชุมชน ในรายวิชาการบริหารสารสนเทศ

The Learning Achievement of the Students Through the Project-based Learning on the Cooperate with the Community on Information Management Course

อภิชาติ เหล็กดี (Apichat Lagdee)

สาขาเทคโนโลยีคอมพิวเตอร์และการสื่อสาร คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏมหาสารคาม
pugan31@gmail.com

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ 1) ศึกษาค่าดัชนีประสิทธิผลของการจัดกิจกรรมการเรียนรู้แบบโครงการร่วมกับชุมชน 2) ศึกษาความพึงพอใจของผู้เรียน ต่อกระบวนการเรียนรู้แบบโครงการร่วมกับชุมชน กลุ่มเป้าหมายที่ใช้ในการวิจัยครั้งนี้ เป็นนักศึกษาชั้นปีที่ 2 คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏมหาสารคาม ภาคเรียนที่ 2 ปีการศึกษา 2556 จำนวน 1 ห้องเรียน มีจำนวนนักศึกษา 31 คน คัดเลือกกลุ่มตัวอย่าง โดยวิธีเจาะจง เครื่องมือที่ใช้ในการวิจัย ได้แก่ แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน และแบบประเมินความพึงพอใจ

ผลการวิจัยสรุปได้ ดังนี้ 1) บทเรียนรายวิชาการบริหารสารสนเทศ โดยใช้กระบวนการเรียนรู้แบบโครงการร่วมกับชุมชน มีประสิทธิภาพเท่ากับ 88.94/81.13 2) ค่าดัชนีประสิทธิผลของการจัดการเรียนรู้ โดยใช้กระบวนการเรียนรู้แบบโครงการร่วมกับชุมชน มีค่าเท่ากับ 0.7936 ซึ่งแสดงว่าผู้เรียนมีผลสัมฤทธิ์ทางการเรียนเพิ่มขึ้น คิดเป็นร้อยละ 79.36 3) นักศึกษาที่เรียนโดยใช้กระบวนการเรียนรู้แบบโครงการร่วมกับชุมชน มีผลสัมฤทธิ์ทางการเรียนหลังเรียนสูงกว่าก่อนเรียน อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และ 4) ความพึงพอใจของนักศึกษาในการเรียนโดยใช้กระบวนการเรียนรู้แบบโครงการร่วมกับชุมชน โดยรวมอยู่ในระดับมากที่สุด มีค่าเฉลี่ยเท่ากับ 4.59 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.08

คำสำคัญ: การเรียนรู้แบบโครงการ โครงการร่วมกับชุมชน
ชุมชน

Abstract

The purposes of this research were 1) Study the effectiveness of the index value 2) Study the student of satisfaction by using the project-based learning on the

cooperate with the community. The target group consisted of 31 , students in the second semester of academic year 2013 from Faculty of Information Technology, Rajabhat Maha Sarakham University. Tools used in the research such as Achievement test and Survey satisfaction.

The research results are as follows:

1) The efficiency of the subjective of education in Information administration by using the project-based learning on the cooperate with the community has criterion of 80/80.

2) The effectiveness of the index value of learning by using the project-based learning on the cooperate with the community has achievement increased 79.36 percent.

3) The student learning achievement attained by using the project-based learning on the cooperate with the community was the average score the posttest was higher than the pretest. The significant relationship was found to be at the 0.01 level. And

4) The student of satisfaction by using the project-based learning on the cooperate with the community has a highest level ($\bar{X} = 4.59$, $SD = 0.08$)

Keywords: The project-based learning, The cooperate with the community, Community

1. บทนำ

การศึกษาเป็นปัจจัยสำคัญในการเรียนรู้ [10] เป็นแหล่งสร้างองค์ความรู้และทรัพยากรมนุษย์ที่สำคัญของสังคม สถาบันอุดมศึกษาเป็นองค์กรพัฒนาการศึกษาระดับสูงที่ไม่

เพียงสูงด้วยคำร่าสาระที่เป็นข้อมูลและรูปแบบเทคโนโลยี หากน่าจะเป็นความสูงด้วยระดับคุณภาพคน โดยเฉพาะอย่างยิ่งสูงด้วยภูมิปัญญา และศีลธรรม มหาวิทยาลัยราชภัฏมหาสารคาม [7] เป็นสถาบันทางการศึกษาที่มีภารกิจตามพระราชบัญญัติ มหาวิทยาลัยราชภัฏ พ.ศ. 2547 มาตรา 6 ที่บัญญัติว่าให้มหาวิทยาลัยเป็นสถาบันอุดมศึกษาเพื่อการพัฒนาท้องถิ่นที่เสริมสร้างพลังปัญญาแห่งแผ่นดิน ฟื้นฟูพลังการเรียนรู้ เชิดชูภูมิปัญญาท้องถิ่น สร้างสรรค์ศิลปวิทยาเพื่อความเจริญก้าวหน้าอย่างมั่นคงและยั่งยืน ซึ่งในแต่ละชุมชนมีภูมิปัญญาท้องถิ่น มีศิลปวัฒนธรรมพื้นบ้านทั้งที่มีการปฏิบัติที่คล้ายคลึงกันและแตกต่างกันออกไปในแต่ละท้องถิ่น ขึ้นอยู่กับบริบทของชุมชน ลักษณะการประกอบอาชีพ และยังขึ้นอยู่กับ การเปิดรับอารยธรรมและความเจริญของกลุ่มชนอื่น เข้ามาผสมผสานระหว่างวัฒนธรรม ซึ่งในปัจจุบันศิลปวัฒนธรรมพื้นบ้านเป็นที่รู้จักในพื้นที่ชุมชนเท่านั้น ไม่มีการเผยแพร่ให้สังคมภายนอกได้รับรู้ และวัฒนธรรมภายนอกเข้ามามีบทบาทหรือแทรกซึมให้วัฒนธรรมพื้นบ้านบางอย่างกลืนหายไปกับกระแส แต่ก็ยังเหลืออีกมากที่ธำรงไว้ซึ่งวัฒนธรรมที่ดีงาม

ในปัจจุบันกระบวนการจัดการเรียนการสอนในระดับอุดมศึกษา [11] มีความหลากหลายและเป็นการจัดการเรียนรู้ที่เน้นผู้เรียนเป็นสำคัญ โดยผ่านการปฏิบัติลงมือ กระทำด้วยตนเอง การพัฒนาศักยภาพ การคิดตลอดจนการเสาะแสวงหาความรู้ด้วยตนเอง ด้วยการปฏิสัมพันธ์กับแหล่งการเรียนรู้ต่างๆ เพื่อนำมาสู่การสร้างความรู้ การเรียนรู้แบบโครงการ [2] เป็นรูปแบบการเรียนการสอนรูปแบบหนึ่งที่ผู้เรียนจะได้ฝึกทักษะรอบด้าน เป็นรูปแบบการเรียนรู้ที่ก่อให้เกิดกระบวนการเรียนรู้ อย่างเป็นระบบ ปลูกฝังความเป็นนักคิดได้อย่างมีประสิทธิภาพ เป็นการเรียนรู้ที่ให้อิสระแก่ผู้เรียน ส่งเสริมความคิดสร้างสรรค์ ในหลากหลายมิติ กิจกรรมการเรียนรู้แบบโครงการเป็นกิจกรรมที่สามารถจัดได้ทั้งในระดับประถมศึกษา มัธยมศึกษา และอุดมศึกษา การเรียนรู้แบบโครงการส่งผลให้ผู้เรียนได้มีโอกาสดำเนินการศึกษา โดยเฉพาะการวางแผนการศึกษา ออกแบบการทดลองหรือวิธีการศึกษา การลงมือทดลองเพื่อตรวจสอบสมมติฐานตลอดจนสรุปผลการศึกษด้วยตนเอง โดยมีผู้ให้คำปรึกษาและชี้แนะ

จากความสำคัญของการจัดการเรียนรู้ที่เน้นผู้เรียนเป็นสำคัญ การเรียนรู้แบบโครงการ บริบทของมหาวิทยาลัยราชภัฏ บริบทของชุมชน และการเปลี่ยนแปลงของเทคโนโลยีสารสนเทศ ผู้วิจัยในฐานะผู้สอนจึงสนใจที่จะจัดกิจกรรมการเรียนรู้ในรายวิชา การบริหารสารสนเทศ โดยใช้กระบวนการเรียนรู้แบบโครงการร่วมกับชุมชน สำหรับนักศึกษาาระดับปริญญาตรี สาขาเทคโนโลยีคอมพิวเตอร์และการสื่อสาร คณะเทคโนโลยีสารสนเทศ เพื่อส่งเสริมการเรียนรู้ให้แก่ผู้เรียน ช่วยให้ผู้เรียนมีความรู้ ความเข้าใจ และมีความสามารถในการจัดทำโครงการร่วมกับชุมชน รวมถึงการมีผลสัมฤทธิ์ทางการเรียนที่ดีขึ้นด้วย โดยเปิดโอกาสให้ผู้เรียนมีส่วนร่วมร่วมกับชุมชน ในการทำโครงการตามความต้องการของชุมชน มีการวางแผนงาน มีส่วนร่วมในการศึกษาแลกเปลี่ยนเรียนรู้ โดยมีผู้วิจัยทำหน้าที่เป็นที่ปรึกษาและฝึกปฏิบัติไปพร้อมกัน ซึ่งจะทำการเรียนการสอนรายวิชาการบริหารสารสนเทศ และการมีส่วนร่วมร่วมกับชุมชน ประสบผลสำเร็จยิ่งขึ้น ผู้เรียนมีความรู้ ความเข้าใจ มีความสามารถในการจัดทำโครงการ ส่งผลให้ผลสัมฤทธิ์ทางการเรียนสูงขึ้นและผู้เรียนมีความพึงพอใจต่อการเรียนมากขึ้น

2. วรรณกรรมที่เกี่ยวข้อง

2.1 Project-based learning

การเรียนรู้แบบโครงการ [9] เป็นการเรียนรู้ที่ยึดผู้เรียนเป็นศูนย์กลาง อาจทำได้หลายวิธี การเรียนรู้แบบให้ผู้เรียนจัดทำโครงการตามความสนใจของผู้เรียน เป็นการเปิดโอกาสให้ผู้เรียนแสวงหาประสบการณ์การเรียนรู้ที่มีความหมายต่อชีวิตประจำวันมากยิ่งขึ้น และสามารถแสดงออกโดยใช้ศักยภาพได้อย่างกว้างขวาง ผู้เรียนจะได้รับการกระตุ้นให้เกิดการเรียนรู้ด้วยการปฏิบัติจริงที่เน้นการสื่อสารและสภาพที่แท้จริง ดังนั้นการกำหนดงานตามความต้องการของผู้เรียน จึงเป็นเป้าหมายหลักสำคัญที่สุด โดยผู้เรียนสามารถเลือกงานที่ต้องการจะทำตามแรงจูงใจของตนเองเป็นหลัก

โครงการ [1] หมายถึง การศึกษาเรื่องใดเรื่องหนึ่งที่นักเรียนเป็นผู้ศึกษา ค้นคว้า และลงมือปฏิบัติด้วยตนเอง ภายใต้คำแนะนำปรึกษาและดูแลของครู อาจารย์ที่ปรึกษาโดยอาจใช้เครื่องมือและอุปกรณ์ต่างๆ ช่วยในการศึกษาค้นคว้านั้น บรรลุผลตามวัตถุประสงค์ การสอนแบบนี้มีประโยชน์มาก มี

ผลในการพัฒนาทักษะ การสร้างเสริมเจตคติ และทำให้นักเรียน มีความพึงพอใจ กระตือรือร้นในการเรียน ให้ความร่วมมือในการปฏิบัติกิจกรรมเป็นอย่างดี

2.2 ศิลปวัฒนธรรมท้องถิ่น

วัฒนธรรมท้องถิ่น หมายถึง การศึกษาเกี่ยวกับสิ่งที่ดีงาม มีคุณค่าซึ่งคนในสังคมประพฤติปฏิบัติหรือแสดงออกมาซ้ำนาน ใน 5 ด้าน คือขนบธรรมเนียมประเพณี และความเชื่อ ภาษาและวรรณกรรม ศิลปกรรมและโบราณคดี การละเล่น ดนตรีและการพักผ่อนหย่อนใจ ชีวิตความเป็นอยู่และวิทยาการ วัฒนธรรมท้องถิ่นเป็นวิถีชีวิตของคนในแต่ละชุมชน ซึ่งเป็นมรดกสืบทอดต่อกันมา และทุกคนในท้องถิ่นย่อมมีความภูมิใจในการเป็นเจ้าของร่วมกัน ทำให้เกิดความรัก ห่วงแหน และศิลปวัฒนธรรมท้องถิ่นเป็นตัวบ่งชี้ความสำคัญถึงวิถีชีวิตของคนในสังคมในแต่ละท้องถิ่นซึ่งสืบทอดกันมา

2.3 การวิจัยปฏิบัติการแบบมีส่วนร่วม (Participatory Action Research : PAR)

การวิจัยปฏิบัติการแบบมีส่วนร่วม หมายถึง การค้นหาความจริงหรือการตอบคำถามที่อยากรู้ โดยมีกระทำการจริงเพื่อบรรลุผลที่ต้องการในการพัฒนา โดยกลุ่มเป้าหมายมีส่วนร่วมในการวิจัยและเรียนรู้แลกเปลี่ยนอย่างสม่ำเสมอ การมีส่วนร่วมเกิดจากแนวคิด 3 ประการ คือ 1) ความสนใจและห่วงกังวลร่วมกัน 2) ความเดือดร้อนและความไม่พึงพอใจร่วมกันต่อสถานการณ์หนึ่งๆ และ 3) การตกลงใจเปลี่ยนแปลงทิศทางที่พึงปรารถนาาร่วมกัน สรุปได้ว่าการวิจัยเชิงปฏิบัติการแบบมีส่วนร่วมเป็นการร่วมคิด ร่วมตัดสินใจ ร่วมเรียนรู้ ร่วมรับผิดชอบ จากสิ่งที่คิดร่วมกันทำร่วมกันตัดสินใจร่วมกันโดยมุ่งเน้นให้เกิดการเปลี่ยนแปลง เกิดการพัฒนา

3. วิธีการดำเนินการวิจัย

3.1 กลุ่มเป้าหมาย เป็นนักศึกษาชั้นปีที่ 2 คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏมหาสารคาม ปีการศึกษา 2556 จำนวน 1 หมู่เรียน จำนวน 31 คน คัดเลือกโดยวิธีเจาะจง

3.2 เครื่องมือที่ใช้ในการวิจัย

- 1) แผนการจัดการเรียนรู้ รายวิชาการบริหารสารสนเทศ ด้วยกระบวนการเรียนรู้แบบโครงงานร่วมกับชุมชน ที่ผ่านการประเมินโดยผู้เชี่ยวชาญ
- 2) แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน รายวิชาการกระบวนการเรียนรู้แบบโครงงานร่วมกับชุมชน ที่ผ่านการประเมินโดยผู้เชี่ยวชาญ
- 3) แบบสอบถามวัดความพึงพอใจที่มีต่อกระบวนการเรียนรู้แบบโครงงานร่วมกับชุมชน ที่ผ่านการประเมินโดยผู้เชี่ยวชาญ

3.3 การดำเนินการทดลอง

3.3.1 การเตรียมการทดลอง

- 1) ผู้วิจัยเตรียมแผนการจัดการเรียนรู้ และเอกสารที่เกี่ยวข้องทั้งหมดตรวจสอบความถูกต้องเรียบร้อย โดยแผนการเรียนรู้ มีดังนี้

ตารางที่ 1 : แผนการเรียนรู้

ลำดับที่	เนื้อหาการเรียนรู้	กิจกรรม
1	หน่วยที่ 1 หลักสำคัญของการบริหารสารสนเทศ	- บรรยาย
2	หน่วยที่ 2 แนวคิด ทฤษฎี หลักการจัดทำโครงงาน	- บรรยาย - แบ่งกลุ่มเพื่อค้นคว้า
3	หน่วยที่ 3 การวิเคราะห์รวบรวมและรักษาข้อมูล และการศึกษาสภาพปัญหา	- บรรยาย - นักศึกษาแสดงความคิดเห็น
4	หน่วยที่ 4 ศึกษาสภาพปัญหาร่วมกับชุมชน	- สอบถาม - ศึกษาสภาพปัญหา
5	หน่วยที่ 5 หลักการเขียนภูมิหลังและสืบค้นทฤษฎี เอกสารงานวิจัยที่เกี่ยวข้อง	- บรรยาย - สาธิตและฝึกปฏิบัติ
6	หน่วยที่ 6 การสืบค้นและการแสดงรายงาน และการเขียนวิธีดำเนินงานโครงการ	- บรรยาย - สาธิตและฝึกปฏิบัติ
7	หน่วยที่ 7 ดำเนินงานโครงการร่วมกับชุมชน	- ศึกษาค้นคว้า - ดำเนินงานร่วมกับชุมชน
8	สอบกลางภาคเรียน	- นำเสนอเค้าโครง
9	หน่วยที่ 8 การเขียนรายงานโครงการ การสรุป อภิปรายผล	- บรรยาย ยกตัวอย่าง - แสดงความคิดเห็น
10 - 11	หน่วยที่ 9 ดำเนินโครงการร่วมกับชุมชน	- ศึกษาค้นคว้า - ดำเนินงานร่วมกับชุมชน

12 - 13	หน่วยที่ 10 สรุปรายงาน ผลการดำเนินงาน ดำเนินโครงการ	- ศึกษาค้นคว้า - เสนอแนะ และให้คำปรึกษา - ดำเนินงานร่วมกับชุมชน
14	การนำเสนอโครงการแบบกลุ่ม	- นำเสนองานเป็นกลุ่ม
15	การนำเสนอโครงการแบบกลุ่ม	- นำเสนองานเป็นกลุ่ม
16	สอบปลายภาคเรียน	สอบภาคทฤษฎี

2) เตรียมสถานที่และเครื่องคอมพิวเตอร์ที่ใช้ในการทดลอง โดยใช้ห้องปฏิบัติการคอมพิวเตอร์ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏมหาสารคาม

3) กำหนดเวลาที่จะทำการทดลอง ในภาคเรียนที่ 2 ปีการศึกษา 2556

3.3.2 ดำเนินการทดลอง ดังนี้

1) ระยะเวลาในการทดลอง ระหว่างวันที่ 28 ตุลาคม 2556 – 28 มกราคม 2557

2) อธิบายวัตถุประสงค์ของรายวิชา เพื่อให้ผู้เรียนมีความรู้ ความเข้าใจ และทักษะเกี่ยวกับการบริหารสารสนเทศ มีทักษะการใช้เทคโนโลยีสารสนเทศและการสื่อสารสมัยใหม่เพื่อการสืบค้นข้อมูล การทำรายงาน การบริหารจัดการสารสนเทศได้อย่างมีประสิทธิภาพ สามารถนำไปประยุกต์ใช้ได้ได้อย่างมีประสิทธิภาพและเหมาะสม และมีจรรยาบรรณในการใช้เทคโนโลยีสารสนเทศอย่างเหมาะสม

2) อธิบายวิธีการเรียน

2.1) ชี้แจงหลักการเรียนรู้ และข้อตกลง

2.2) อธิบายความรู้เบื้องต้นเกี่ยวกับการทำโครงการเบื้องต้น

2.3) อธิบายวิธีการจัดทำโครงการร่วมกับชุมชน และวิธีการเข้าถึงชุมชน

2.4) ผู้เรียนศึกษารายละเอียดเพิ่มเติมจากใบความรู้ที่จัดทำไว้ในแผนการจัดการเรียนรู้ในแต่ละแผน

2.5) ผู้เรียนปฏิบัติกิจกรรมการจัดทำโครงการตามที่กำหนดไว้ในแต่ละแผนการจัดการเรียนรู้

3) ทำการทดลอง ผู้วิจัยดำเนินการทดลองจัดการเรียนรู้โดยใช้กระบวนการเรียนรู้แบบโครงการร่วมกับชุมชน โดยใช้แผนการจัดการเรียนรู้แบบโครงการ พร้อมทั้งทำแบบทดสอบก่อนเรียน แบบฝึกปฏิบัติ ใบงาน และ

แบบทดสอบหลังเรียน เพื่อทำการวิเคราะห์และประสิทธิภาพของกระบวนการต่อไป ดังนี้

3.1) แจกจุดประสงค์ในการเรียนให้กลุ่มทดลองทราบ

3.2) นำแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนที่หาคุณภาพแล้วมาทดสอบกับนักเรียนก่อนเรียน (Pre-test)

3.3) ดำเนินการสอนโดยใช้เวลา ตามวันและเวลาในตารางเรียนปกติ รวมถึงการมีส่วนร่วมกับชุมชน

3.4) นำแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนที่หาคุณภาพแล้วมาทดสอบกับนักเรียนหลังเรียนเสร็จทุกบทเรียนแล้ว (Pos-test)

3.5) ให้นักเรียนในกลุ่มทดลองทำแบบประเมินวัดความพึงพอใจหลังจากแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนหลังเรียนเสร็จสิ้นแล้ว

3.6) รวบรวมข้อมูลทั้งหมดและวิเคราะห์ข้อมูลโดยใช้วิธีการทางสถิติ

3.7) สรุปผลการทดลอง

3.4 การวิเคราะห์ข้อมูล

1) การวิเคราะห์หาประสิทธิภาพของบทเรียนรายวิชาการบริหารสารสนเทศ โดยใช้กระบวนการเรียนรู้แบบโครงการร่วมกับชุมชน จากค่า E_1/E_2

2) การวิเคราะห์หาค่าดัชนีประสิทธิผลของกระบวนการเรียนรู้แบบโครงการร่วมกับชุมชน โดยใช้สูตรการหาค่าดัชนีประสิทธิผล (Effectiveness Index ; E.I.) ของกู๊ดแมน, เฟลทเชอร์และชไนเดอร์ [5]

3) วิเคราะห์เปรียบเทียบผลสัมฤทธิ์ทางการเรียนของนักเรียนระหว่างการสอนด้วยกระบวนการเรียนรู้แบบโครงการร่วมกับชุมชนกับการสอนด้วยวิธีปกติ โดยใช้สถิติ t-test (Dependent Samples)

4) การวิเคราะห์หาความพึงพอใจต่อกระบวนการเรียนรู้แบบโครงการร่วมกับชุมชน โดยใช้ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน [4]

4. ผลการดำเนินงาน

ตอนที่ 1 วิเคราะห์หาประสิทธิภาพของบทเรียนรายวิชาการบริหารสารสนเทศ โดยใช้กระบวนการเรียนรู้แบบ

โครงการร่วมกับชุมชน ดำเนินการหาประสิทธิภาพตามเกณฑ์ 80/80 ดังนี้

ตารางที่ 2 : แสดงผลการวิเคราะห์หาประสิทธิภาพ

ผลการทดสอบ	N	คะแนน เต็ม	คะแนน รวม	\bar{X}	ร้อยละ
ระหว่างการ ทดลอง	31	20	439	16.26	88.94
หลังทดลอง	31	10	231	8.56	81.13

จากตารางที่ 2 ผลการทดลองพบว่า ผลการทดสอบระหว่างทดลองของนักศึกษาที่เรียนด้วยบทเรียนรายวิชา การบริหารสารสนเทศ โดยใช้กระบวนการเรียนรู้แบบโครงการร่วมกับชุมชน มีค่าเฉลี่ยเท่ากับ 16.26 คิดเป็นร้อยละ 88.94 และผลการทดลองหลังเรียน มีค่าเฉลี่ยเท่ากับ 8.56 คิดเป็นร้อยละ 81.13 ดังนั้น บทเรียนรายวิชาการบริหารสารสนเทศ โดยใช้กระบวนการเรียนรู้แบบโครงการร่วมกับชุมชน มีประสิทธิภาพของกระบวนการ (E_1) เท่ากับ 88.94 และประสิทธิภาพของผลลัพธ์ (E_2) เท่ากับ 81.13 ดังนั้นมีประสิทธิภาพเท่ากับ 88.94 / 81.13 ซึ่งเป็นไปตามเกณฑ์ที่กำหนดไว้

ตอนที่ 2 วิเคราะห์หาค่าดัชนีประสิทธิผลของการจัดการเรียนรู้โดยใช้กระบวนการเรียนรู้แบบโครงการร่วมกับชุมชน ได้ผลการวิเคราะห์ข้อมูล ดังนี้

ตารางที่ 3 : แสดงค่าดัชนีประสิทธิผลของกระบวนการเรียนรู้

N	คะแนน เต็ม	ผลรวมของคะแนน		ร้อยละ		E.I.
		ก่อนเรียน	หลัง เรียน	ก่อน เรียน	หลัง เรียน	
31	10	81.00	231.00	30.00	81.13	0.7936

จากตารางที่ 3 พบว่า ผลการวิเคราะห์ค่าดัชนีประสิทธิผลของการเรียนรู้โดยใช้กระบวนการเรียนรู้แบบโครงการร่วมกับชุมชน มีค่าเท่ากับ 0.7936 หมายความว่าผู้เรียนมีคะแนนเพิ่มขึ้นหลังจากการเรียนรู้ด้วยกระบวนการเรียนรู้แบบโครงการร่วมกับชุมชน คิดเป็นร้อยละ 79.36

ตอนที่ 3 ผลการเปรียบเทียบผลสัมฤทธิ์ทางการเรียนของนักศึกษาหลังเรียนโดยใช้กระบวนการเรียนรู้แบบโครงการร่วมกับชุมชน สำหรับนักศึกษาชั้นปีที่ 2 คณะเทคโนโลยี

สารสนเทศ มหาวิทยาลัยราชภัฏมหาสารคาม กับคะแนนก่อนเรียน ดังนี้

ตารางที่ 4 : แสดงผลวิเคราะห์เปรียบเทียบผลสัมฤทธิ์ทางการเรียน

ผลสัมฤทธิ์	N	\bar{X}	S.D.	$\sum D$	$\sum D^2$	$(\sum D)^2$	t-test
ก่อนเรียน	31	3.00	0.87	179	1,787	32,041	17368**
หลังเรียน	31	8.56	2.68				

01. มีนัยสำคัญทางสถิติที่ระดับ ** ($t_{0.01, 18} = 2.552$) , df = 18 จากตารางที่ 4 พบว่า ผลสัมฤทธิ์ทางการเรียนของนักศึกษาหลังเรียน โดยใช้กระบวนการเรียนรู้แบบโครงการร่วมกับชุมชน กับคะแนนก่อนเรียนที่ผู้วิจัยพัฒนาขึ้นสูงกว่าก่อนเรียนอย่างมีนัยสำคัญที่ระดับ 0.01

ตอนที่ 4 ศึกษาความพึงพอใจของนักศึกษาในการเรียน โดยใช้กระบวนการเรียนรู้แบบโครงการร่วมกับชุมชน ได้ผลการวิเคราะห์ข้อมูล ดังนี้

ตารางที่ 5 : แสดงผลการวิเคราะห์ข้อมูลจากการสอบถามความพึงพอใจของนักศึกษา

ข้อ	ข้อความ	\bar{X}	S.D.	ระดับความ พึงพอใจ
1	ผู้เรียนมีความเข้าใจในเนื้อหาการเรียนจากบทเรียน แผน การเรียนรู้ ใบความรู้ ใบงาน	4.65	0.49	มากที่สุด
2	ผู้เรียนสามารถเรียนรู้เนื้อหาจากบทเรียน ได้สะดวกและรวดเร็ว	4.55	0.60	มากที่สุด
3	เนื้อหาของบทเรียน แบบฝึก ใบความรู้ เหมาะสม	4.35	0.67	มาก
4	เนื้อหาของบทเรียน แบบฝึกปฏิบัติมีความน่าสนใจ	4.60	0.60	มากที่สุด
5	ผู้เรียนสามารถอ่านและทำความเข้าใจในเนื้อหาได้ด้วยตนเองสามารถทำโครงการได้อย่างมีประสิทธิภาพ	4.55	0.51	มากที่สุด
6	เนื้อหาของบทเรียนที่ได้ฝึกปฏิบัติร่วมกับชุมชน ทำให้การเรียนมีความน่าสนใจ	4.50	0.61	มากที่สุด
7	ผู้เรียน มีความเอาใจใส่ต่อการเรียนมากขึ้นทำให้เกิดความรู้จากการปฏิบัติและเรียนรู้ด้วยตนเองมากขึ้น	4.85	0.37	มากที่สุด
8	การทำโครงการร่วมกับชุมชนให้ทั้งความรู้ ทักษะ และประสบการณ์	4.65	0.59	มากที่สุด
9	การฝึกปฏิบัติร่วมกับชุมชนช่วยแก้	4.55	0.60	มากที่สุด

ข้อ	ข้อความ	\bar{X}	SD.	ระดับความพึงพอใจ
	การเรียนรู้ที่ไม่เข้าใจได้			
10	การฝึกปฏิบัติร่วมกับชุมชน ช่วยให้เข้าใจภูมิปัญญาท้องถิ่น	4.65	0.67	มากที่สุด
11	การฝึกปฏิบัติร่วมกับชุมชนช่วยเพิ่มทักษะในการแสวงหาความรู้	4.65	0.59	มากที่สุด
12	การฝึกปฏิบัติร่วมกับชุมชนช่วยเพิ่มทักษะในการแก้ปัญหาอย่างเป็นระบบ	4.65	0.49	มากที่สุด
13	การฝึกทักษะการมีส่วนร่วมร่วมกับชุมชนช่วยเพิ่มทักษะและนำไปใช้ประโยชน์ในชีวิตประจำวัน และการประกอบอาชีพ	4.60	0.68	มากที่สุด
14	นักเรียนได้ฝึกปฏิบัติจริง	4.40	0.60	มาก
15	การเตรียมความพร้อมด้านสื่อ และวัสดุอุปกรณ์ในการเรียน	4.60	0.60	มากที่สุด
	รวมเฉลี่ย	4.59	0.08	มากที่สุด

จากตารางที่ 5 พบว่า ความพึงพอใจของนักศึกษาในการเรียนรู้โดยใช้กระบวนการเรียนรู้แบบโครงการร่วมกับชุมชนโดยรวมอยู่ในระดับมากที่สุด มีค่าเฉลี่ยเท่ากับ 4.59 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.08 เมื่อพิจารณาเป็นรายข้อ พบว่าข้อที่มีค่าเฉลี่ยมากที่สุด คือ ผู้เรียนมีความเอาใจใส่ต่อการเรียนมากขึ้นทำให้เกิดความรู้จักการปฏิบัติและเรียนรู้ด้วยตนเองมากขึ้น มีค่าเฉลี่ยเท่ากับ 4.85 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.37 และ ข้อที่มีค่าเฉลี่ยน้อยที่สุด คือ เนื้อหาของบทเรียนแบบฝึกหัด ใบความรู้ เหมาะสม มีค่าเฉลี่ยเท่ากับ 4.35 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.67

5. สรุปผลการวิจัย

จากผลการทดลองสรุปว่า การจัดการเรียนรู้ด้วยกระบวนการเรียนรู้แบบโครงการร่วมกับชุมชน มีประสิทธิภาพตามเกณฑ์ที่ตั้ง เท่ากับ 88.94/81.13 มีค่าดัชนีประสิทธิผลของการจัดการเรียนรู้ เท่ากับ 0.7936 ซึ่งแสดงว่านักศึกษามีความก้าวหน้าทางการเรียนเพิ่มขึ้นคิดเป็นร้อยละ 79.36 นักศึกษาที่เรียนโดยกระบวนการเรียนรู้แบบโครงการร่วมกับชุมชน มีผลสัมฤทธิ์ทางการเรียนหลังเรียนสูงกว่าก่อนเรียนอย่างมีนัยสำคัญทางสถิติที่ ระดับ .01 ความพึงพอใจของนักศึกษาในการเรียนโดยกระบวนการเรียนรู้แบบ

โครงการร่วมกับชุมชน โดยรวมอยู่ในระดับมากที่สุด มีค่าเฉลี่ยเท่ากับ 4.59 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.08 ดังนั้น กระบวนการเรียนรู้แบบโครงการร่วมกับชุมชนที่ผู้วิจัยพัฒนาขึ้น เป็นกระบวนการที่มีประสิทธิภาพ สามารถนำไปใช้ในการจัดการเรียนรู้ในรายวิชาอื่นได้

เอกสารอ้างอิง

- [1] กรมวิชาการ กระทรวงศึกษาธิการ. เอกสารชุดเทคนิคการจัดกระบวนการเรียนรู้ที่ผู้เรียนสำคัญที่สุด. กรุงเทพฯ : ศูนย์พัฒนาหลักสูตร กรมวิชาการ, 2544.
- [2] ถวัลย์ มาศจรัส และ มณี เรืองขำ. แนวการจัดการกิจกรรมการเรียนการสอนโครงการ. กรุงเทพฯ : ชารักษ์, 2549.
- [3] ทิศนา แจมมณี. รูปแบบการสอน : ทางเลือกที่หลากหลาย. กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2551.
- [4] บุญชม ศรีสะอาด. การวิจัยเบื้องต้น. กรุงเทพฯ, 2545.
- [5] เผชญิ กิจระการ. ดัชนีประสิทธิผล Effectiveness Index. มหาสารคาม : คณะศึกษาศาสตร์ มหาวิทยาลัยราชภัฏมหาสารคาม, 2544.
- [6] พิมพ์ เคชะคุปต์. การสอนคิดด้วยโครงการ : การเรียนการสอนบูรณาการ. กรุงเทพฯ : โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย, 2551.
- [7] มหาวิทยาลัยราชภัฏมหาสารคาม. แผนยุทธศาสตร์มหาวิทยาลัยราชภัฏมหาสารคาม ฉบับที่ 11 (พ.ศ. 2555-2561). มหาสารคาม : มหาวิทยาลัยราชภัฏมหาสารคาม, 2555.
- [8] วราภรณ์ ตรีคุณสถิตย์. แนวคิดการจัดการเรียนรู้แบบโครงการ. กรุงเทพฯ : เอ็ม ไอ ที พรินต์ติ้ง, 2551.
- [9] สมศักดิ์ ภู่วิภาดาพรรณ. เทคนิคการส่งเสริมความคิดสร้างสรรค์. กรุงเทพฯ : โรงพิมพ์ไทยวัฒนาพานิช, 2544.
- [10] สายชล จินใจการพัฒนา รูปแบบการเรียนการสอนแลด . ผสมผสานรายวิชาการเขียนโปรแกรมคอมพิวเตอร์ 1 สาขาคอมพิวเตอร์ธุรกิจปริญญาโทปริญญาตรี . บัณฑิต สาขาคอมพิวเตอร์ศึกษา. กรุงเทพฯ : มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2550.
- [11] สุมาลี ชัยเจริญ และอิสรา ก้านจักร. ปรัชญาแห่งความรู้ (Epistemology) : จากปรัชญาสู่การปฏิบัติ. วารสารเทคโนโลยีทางปัญญา. ขอนแก่น : มหาวิทยาลัยขอนแก่น, 2550.

การพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอน เรื่อง การอ่านค่าอัตราส่วนแนวตั้งระหว่าง Cup และ Disc (CDR) จากภาพถ่ายจอประสาทตาสำหรับนักศึกษาแพทย์เฉพาะทางจักษุวิทยา
A Development of Computer-Assisted Instruction Program on the Topic of “Estimation of Vertical Cup-to-Disc Ratio from Retinal Fundus Images” for Eye Residents

นฤมล นวลฉวี (Narumol Nuanchawe)¹ และ นงลักษณ์ โควาวีสารัช (Nongluk Covavisaruch)²

ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

¹Narumol.nua@chula.ac.th, ²Nongluk.c@chula.ac.th

บทคัดย่อ

บทเรียนคอมพิวเตอร์ช่วยสอนเรื่อง การอ่านค่าอัตราส่วนแนวตั้งระหว่าง Cup และ Disc (CDR) จากภาพถ่ายจอประสาทตา Retinal fundus image นี้พัฒนาโดยใช้เทคนิคการใช้ภาพเป็นตัวประสานกับผู้ใช้ Graphical User Interface (GUI) เพื่อให้ช่วยให้ผลสัมฤทธิ์ทางการเรียนรู้สูงขึ้นและสร้างความพึงพอใจต่อบทเรียนช่วยสอนที่สร้างความแม่นยำในการวินิจฉัยโรคต้อหินเพิ่มมากขึ้น

งานวิจัยนี้ทดสอบการใช้งานกับกลุ่มตัวอย่างที่เป็นนักศึกษาแพทย์เฉพาะทางด้านจักษุวิทยา คณะแพทยศาสตร์โรงพยาบาลรามาธิบดี จำนวน 15 คน เครื่องมือที่ใช้คือ โปรแกรมแบบทดสอบวัดผลสัมฤทธิ์ในการอ่านค่าอัตราส่วน CDR จากภาพถ่ายจอประสาทตาจำนวน 10 ข้อ และ แบบประเมินความพึงพอใจ สถิติที่ใช้ในการวิเคราะห์ข้อมูล ได้แก่ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ทดสอบสมมติฐาน ด้วย t-test แบบ Dependent Sample

ได้ผลสัมฤทธิ์ของผู้เรียนก่อนและหลังได้รับการเรียนรู้ด้วยบทเรียนคอมพิวเตอร์ช่วยสอนที่พัฒนาขึ้นสูงกว่าก่อนเรียนอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 โดยคะแนนเฉลี่ยหลังเรียน (\bar{X} = 8.33) ซึ่งสูงกว่าคะแนนเฉลี่ยก่อนเรียน (\bar{X} = 4.20) และพบว่าหลังจากได้ผ่านการใช้บทเรียนคอมพิวเตอร์ช่วยสอนนี้ ผู้เรียนมีความพึงพอใจต่อบทเรียนโดยรวมอยู่ในระดับมากที่สุด (\bar{X} = 4.55, S.D. = 0.61)

จากการวิจัยสรุปได้ว่า การพัฒนาบทเรียนช่วยสอนมีการใช้เทคนิคการใช้ภาพเป็นตัวประสานกับผู้ใช้ Graphical User

Interface (GUI) ซึ่งผลการพัฒนาระบบออกมาอยู่ในระดับดีมาก รายละเอียดอธิบายในบทที่ 4 และโปรแกรมสามารถนำไปประยุกต์ใช้ได้จริง

คำสำคัญ: คอมพิวเตอร์ช่วยสอน CDR ภาพถ่ายจอประสาทตา

Abstract

A GUI-based (Graphical User Interface) CAI (computer-assisted instruction) program on the topic of “Estimation of vertical cup-to-disc ratio from retinal fundus images” was developed to help improve the efficiency of eye residents’ study and to satisfy users.

In this study, users are the eye resident from Ramathibodi hospital. User consists of 15 The test has 10 images. We apply the statistic to prove the user satisfaction about our application such as Mean, Standard deviation (SD) and T-test type Dependent Sample.

The result shows as the following: first, the score after user pass our CAI is higher. The Level of Significance is 0.05 and Mean value after practice is 8.33 which increasing from 4.20 (before using our CAI. Second, users are satisfy with our CAI too much (\bar{X} = 4.55 and S.D. = 0.61)

Finally, we apply the practice using Computer Graphical User Interface (GUI) lead to higher score which detail will describe in chapter 4. Our application

shows, it can be a tool for practice the Glaucoma Disease to the clinician.

Keywords: Computer-assisted Instruction ,CDR,Retinal Fundus Image

1. บทนำ

ในการเรียนการสอนของนักศึกษาแพทย์จะเป็นการเรียนทฤษฎีและฝึกปฏิบัติการตรวจกับหุ่นและผู้ป่วยจริง โดยมีอาจารย์แพทย์ผู้เชี่ยวชาญเฉพาะด้านเป็นผู้กำกับดูแลเพื่อให้การตรวจโรคถูกต้องตามขั้นตอนและเพื่อให้อาจารย์แพทย์ช่วยยืนยันการตัดสินใจวินิจฉัยโรคและการให้การรักษา อย่างไรก็ตาม ปัญหาที่พบเป็นประจำในการเรียนการสอนนักศึกษาแพทย์มีสองอย่างหลัก ๆ คือ การไม่ได้รับความร่วมมือจากผู้ป่วยจริงในคลินิกและการที่มีอาจารย์แพทย์ผู้เชี่ยวชาญจำนวนน้อยซึ่งทำให้การยืนยันผลการตรวจจากอาจารย์แพทย์ต้องใช้เวลา

ในการศึกษาเฉพาะทางของจักษุแพทย์เพื่อการวินิจฉัยโรคต้อหิน (Glaucoma) มีหลักการตรวจวินิจฉัยโดยการตรวจจอประสาทตาแล้วอ่านค่าอัตราส่วนแนวตั้งระหว่าง Cup และ Disc (Vertical cup-to-disc ratio ซึ่งต่อไปนี้จะเรียกโดยย่อว่า CDR) จากภาพถ่ายจอประสาทตา (Retinal fundus image) ประกอบกับข้อมูลที่ได้จากการซักถามประวัติและอาการของผู้ป่วย การอ่านค่า CDR จึงเป็นทักษะที่สำคัญและจำเป็นสำหรับจักษุแพทย์ด้านโรคต้อหิน

ในทางปฏิบัติ การฝึกฝนทักษะการอ่านค่า CDR ของนักศึกษาแพทย์ขึ้นกับหลายปัจจัย กล่าวคือ จำนวนผู้ป่วยที่นักศึกษาแพทย์มีโอกาสได้ตรวจ ความหลากหลายของลักษณะต้อหินของผู้ป่วย ข้อจำกัดในแง่ของเวลาในการตรวจ และการไม่มีโอกาสที่จะฝึกทักษะการอ่านผลหลาย ๆ ครั้งเพื่อเพิ่มความแม่นยำในการอ่านผลให้กับตนเอง

ด้วยปัจจุบัน เทคโนโลยีทางคอมพิวเตอร์มีบทบาทโดยตรงกับระบบการศึกษาในการนำเสนอและการแสดงผลด้วยระบบสื่อต่าง ๆ ทั้งในด้านข้อมูล รูปภาพ เสียง ภาพเคลื่อนไหว และวิดีโอ นอกจากนี้ ยังสามารถสร้างระบบการมีปฏิสัมพันธ์แบบโต้ตอบระหว่างคอมพิวเตอร์กับผู้ใช้งาน ทำให้การเรียนรู้ยุคใหม่ประสบความสำเร็จด้วยดี [1] นักการศึกษาและนักวิจัยนำ

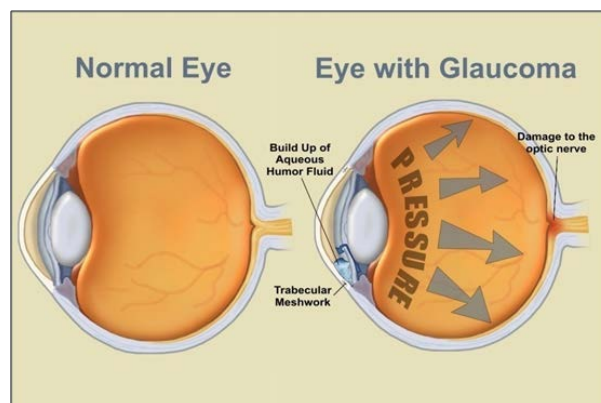
เทคโนโลยีคอมพิวเตอร์มาใช้ในบทบาทเป็นผู้สอนอย่างกว้างขวาง ทั้งในประเทศและต่างประเทศ [2] พบว่า การเรียนโดยใช้คอมพิวเตอร์มีบทบาทเป็นผู้สอนทำให้ผลสัมฤทธิ์ทางการเรียนของผู้เรียนสูงขึ้น เมื่อเทียบกับวิธีสอนแบบปกติที่ใช้ครูเป็นผู้สอน

จากปัญหาข้างต้น งานวิจัยนี้จึงมีแนวคิดที่จะพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนเพื่อ 1) ใช้ประกอบการจัดการเรียนรู้เรื่องการอ่านค่าอัตราส่วนแนวตั้งระหว่าง Cup และ Disc (CDR) จากภาพถ่ายจอประสาทตา (Retinal fundus image) ของนักศึกษาแพทย์เฉพาะทางด้านจักษุวิทยา คณะแพทยศาสตร์โรงพยาบาลรามาธิบดี 2) บทเรียนช่วยสอนนี้จะช่วยให้ผลสัมฤทธิ์ทางการเรียนสูงขึ้น และ 3) บทเรียนช่วยสอนนี้จะสร้างความพึงพอใจต่อบทเรียนช่วยสอนที่สร้างความแม่นยำในการวินิจฉัยโรคต้อหินเพิ่มมากขึ้น

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 โรคต้อหิน

โรคต้อหินเป็นโรคที่ผู้ป่วยจะค่อยๆ สูญเสียการมองเห็น มักเกิดพร้อมกันทั้งสองข้างและไม่แสดงอาการใดๆ และผู้ป่วยไม่ทราบว่าตนเองเป็นโรคต้อหิน



ภาพที่ 1: แสดงลูกต้อปกติและลูกต้อที่เป็นโรคต้อหิน

จากภาพที่ 1 ภาพซ้ายมือเป็นภาพของต้อปกติ ส่วนภาพขวามืออธิบายการเกิดโรคต้อหิน โรคต้อหินเกิดจากระบบหมุนเวียนของของเหลวภายในลูกตามีความผิดปกติ ทำให้มีการเพิ่มขึ้นของของเหลว ส่งผลให้ความดันภายในลูกตาสูงขึ้น ทำให้เส้นประสาทตา (Optic nerve) ซึ่งเป็นเส้นประสาทที่มีความสำคัญต่อการมองเห็นถูกทำลายจากความดันภายในลูกตาที่เพิ่มขึ้น [4]

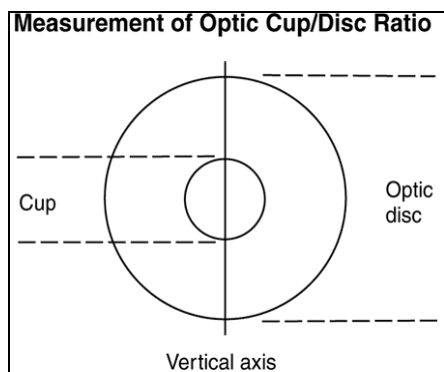
2.2 การวินิจฉัยโรคต้อหิน

ขั้นตอนโดยทั่วไปในการวินิจฉัยโรคต้อหิน คือ

- 1) ซักประวัติทางร่างกายและประวัติทางครอบครัว
- 2) วัดสายตาเพื่อวัดการมองเห็นว่าอยู่ในเกณฑ์ปกติหรือไม่
- 3) วัดความดันภายในลูกตาเพราะความดันภายในลูกตาเป็นปัจจัยในการเกิดโรคต้อหิน
- 4) ตรวจดูขั้วประสาทตาและจอตา โดยแพทย์จะตรวจรูปร่างลักษณะของขั้วประสาทตาจากภาพถ่ายจอประสาทตา ดังภาพที่ 2 ซึ่งได้จากกล้องถ่ายภาพจอประสาทตา (Retinal fundus camera) เพื่ออ่านค่า CDR ดังภาพที่ 3



ภาพที่ 2: แสดงภาพถ่ายจอประสาทตาที่ถ่ายด้วยกล้องถ่ายภาพจอประสาทตา [4]



ภาพที่ 3: การอ่านค่าอัตราส่วนแนวตั้งระหว่าง Cup และ Disc (CDR) จากภาพถ่ายจอประสาทตา [5]

2.2 งานวิจัยที่เกี่ยวข้อง

ทิพย์มณฑา สดชื่น ได้พัฒนาหนังสืออิเล็กทรอนิกส์เรื่อง การถ่ายภาพเบื้องต้น สำหรับนักศึกษาคณะศึกษาศาสตร์ที่ไม่เคยเรียนวิชา 263-201 เทคโนโลยีทางการศึกษา มาก่อน โดยมีวัตถุประสงค์เพื่อหาประสิทธิภาพหนังสืออิเล็กทรอนิกส์ตามเกณฑ์มาตรฐาน 80/80 และเพื่อเปรียบเทียบผลสัมฤทธิ์ทางการเรียน ก่อนเรียนและหลังเรียนของนักศึกษาที่เรียน โดยใช้

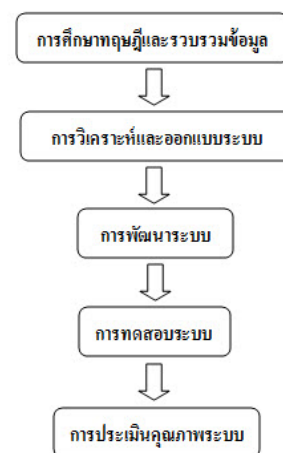
หนังสืออิเล็กทรอนิกส์นี้ ผลการวิจัยพบว่า หนังสืออิเล็กทรอนิกส์เรื่อง การถ่ายภาพเบื้องต้น มีประสิทธิภาพ 87.5/83.44 ซึ่งสูงกว่าเกณฑ์มาตรฐาน 80/80 ส่วนผลสัมฤทธิ์ทางการเรียนของนักศึกษาหลังจากที่เรียนด้วยหนังสืออิเล็กทรอนิกส์นี้สูงกว่าก่อนเรียนอย่างมีนัยสำคัญทางสถิติ ที่ระดับ .01 [6]

อินทิรา ชูศรีทอง ได้ศึกษาวิจัยการพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนระบบมัลติมีเดีย วิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 ผลการวิจัยพบว่า บทเรียนคอมพิวเตอร์ช่วยสอนมีประสิทธิภาพ 86/88 ความคงทนในการเรียนรู้ทดสอบวัดผลสัมฤทธิ์ทางการเรียนหลังจากนักเรียนได้เรียนด้วยบทเรียนคอมพิวเตอร์ช่วยสอนผ่านไปแล้ว 15 วัน ปรากฏว่าคะแนนเฉลี่ยลดลงร้อยละ 2.28 และผ่านไป 30 วันคะแนนเฉลี่ยลดลงร้อยละ 3.22 และความคิดเห็นของนักเรียนที่มีต่อการเรียนด้วยบทเรียนคอมพิวเตอร์ช่วยสอนโดยรวมเห็นด้วยระดับมาก [7]

จากงานวิจัยที่ผ่านมาได้มีการนำภาพถ่ายและระบบมัลติมีเดียมาใช้เป็นบทเรียนคอมพิวเตอร์ช่วยสอนมากขึ้น จึงมีได้มีแนวคิดในการพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอน เรื่อง การอ่านค่าอัตราส่วนแนวตั้งระหว่าง Cup และ Disc (CDR) จากภาพถ่ายจอประสาท Retinal fundus image

3. วิธีการดำเนินการวิจัย

ผู้วิจัยได้ออกแบบขั้นตอนการพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอน เรื่อง การอ่านค่าอัตราส่วนแนวตั้งจากภาพถ่ายจอประสาทตา โดยแบ่งออกเป็น 5 ขั้นตอนดังแสดงในภาพที่ 4 ดังนี้



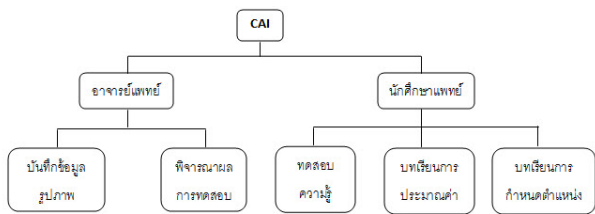
ภาพที่ 4: ขั้นตอนการพัฒนาบทเรียน

3.1 การศึกษาทฤษฎีและรวบรวมข้อมูล

จากการศึกษาวิธีการเรียนการสอนการวินิจฉัยโรคต้อหิน ของนักศึกษาแพทย์เฉพาะทางจักษุวิทยาและพบปัญหาที่สามารถนำเทคโนโลยีคอมพิวเตอร์ที่ใช้ช่วยสอน ผู้วิจัยจึงได้หาข้อมูลภาพถ่ายจอประสาทตาที่ช่วยในการวินิจฉัยโรคต้อหินได้ พบว่ามีเก็บข้อมูลภาพถ่ายจอประสาทตาของผู้ป่วยที่ได้เข้ารับการรักษาเป็นจำนวนมาก แต่ก็ยังไม่มีการนำข้อมูลรูปภาพดังกล่าวมาทำประโยชน์เพื่อให้เกิดการเรียนรู้กับนักศึกษาแพทย์

3.2 การวิเคราะห์และออกแบบระบบ

จากการวิเคราะห์ข้อมูลและลักษณะของงาน พบว่ามีความต้องการให้ผู้เรียนเห็นภาพและมีคำถามและตอบคำถาม ดังนั้นเทคนิคการใช้ภาพเป็นตัวประสานกับผู้ใช้ Graphical User Interface (GUI) จึงเป็นวิธีการที่เหมาะสมกับงานที่ให้ความสะดวกแก่ ผู้ใช้คอมพิวเตอร์ ให้ติดต่อสื่อสารกับเครื่องคอมพิวเตอร์โดยผ่านทางภาพ



ภาพที่ 5: โครงสร้างของระบบโดยรวม

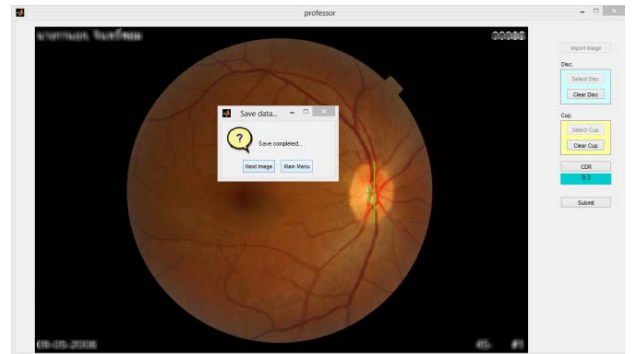
ภาพที่ 5 แสดงโครงสร้างของระบบช่วยสอนซึ่งแบ่งเป็น 2 ส่วน คือ ส่วนของอาจารย์แพทย์และส่วนของนักศึกษาแพทย์ ส่วนแรกเป็นส่วนที่อาจารย์แพทย์ที่เป็นผู้สร้างบทเรียน นำรูปภาพที่ใช้บันทึกเป็นบทเรียนและเป็นข้อทดสอบพร้อมเฉลยคำตอบเข้าระบบ รวมทั้งสามารถติดตาม ผลการทดสอบของนักศึกษาแพทย์ ผ่านการบันทึกข้อมูลผลการทดสอบผ่านโปรแกรม Excel สำหรับส่วนที่สองเป็นส่วนที่นักศึกษาแพทย์ใช้ศึกษาบทเรียนช่วยสอนและทำข้อทดสอบก่อนและหลังเรียน

3.3 การพัฒนาระบบ

ผู้วิจัยพัฒนา โปรแกรม ใน ลักษณะ ของ Windows Application โดยการใช้ภาพประสานงานกับผู้ใช้ GUI ด้วยภาษาซี มีการใช้งานแบ่งเป็น 2 ส่วน ดังนี้

ส่วนใช้งานของอาจารย์แพทย์ มีหน้าจอการใช้งานสำหรับการนำเข้าภาพ แล้วกำหนดตำแหน่งส่วนสูงแนวตั้งของ Cup และ Disc ในภาพถ่ายจอประสาทตาดังแสดงในภาพที่ 6 เป็นเส้นตรง

แนวตั้งสองเส้น โดยเส้นสั้น คือส่วนสูงของ Cup และเส้นยาวสำหรับ Disc จากนั้นมีการแสดงค่า CDR ที่คำนวณได้แล้ว บันทึกข้อมูลและคูณผลผ่าน โปรแกรม Excel ดังภาพที่ 7 โดยมีการเก็บข้อมูลเป็นตำแหน่งจุดปลายของเส้นกำหนดส่วนสูงของ Disc และ Cup ที่อาจารย์แพทย์ให้ไว้ ข้อมูลค่า CDR และชื่อเพิ่มข้อมูลภาพ



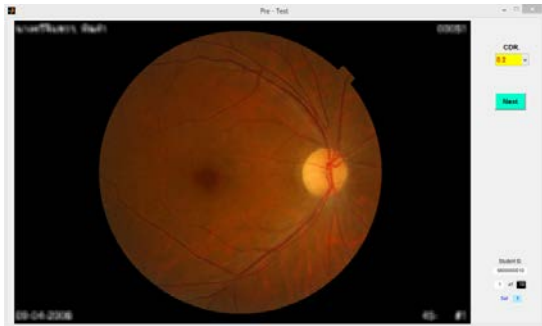
ภาพที่ 6: หน้าจอแสดงส่วนการใช้งานของอาจารย์แพทย์

	A	B	C	D	E	F	G	H	I	J
1	disc_x1	disc_y1	disc_x2	disc_y2	cup_x1	cup_y1	cup_x2	cup_y2	cdr	filename
2	2057	801	2054	1160	2017	878	2017	1055	0.5	E0101.jpg
3	912	792	912	1104	967	866	967	1011	0.5	E0102.jpg
4	998	866	1001	1126	1051	943	1054	1033	0.3	E0103.jpg
5	2063	906	2063	1231	2011	971	2017	1154	0.6	E0104.jpg
6	2073	826	2070	1206	2035	971	2039	1076	0.3	E0105.jpg
7	2048	869	2051	1120	1998	925	1995	1048	0.5	E0106.jpg
8	2048	835	2045	1179	2014	931	2014	1089	0.5	E0107.jpg
9	1918	603	1915	1135	1983	742	1983	974	0.4	E0108.jpg
10	1952	761	1952	1278	2011	887	2011	1114	0.4	E0109.jpg
11	1933	860	1933	1160	1884	940	1884	1061	0.4	E0110.jpg

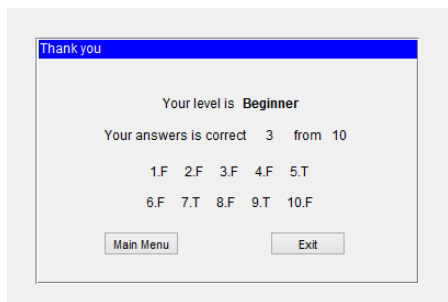
ภาพที่ 7: แสดงผลการบันทึกข้อมูลผ่าน โปรแกรม Excel

ส่วนใช้งานของนักศึกษาแพทย์ มีหน้าจอการใช้งาน 3 ส่วน คือ “หน้าจอสำหรับการทดสอบความรู้ก่อนและหลังเรียน” “หน้าจอสำหรับฝึกทักษะการประมาณค่า CDR” และ “หน้าจอสำหรับฝึกทักษะการกำหนดตำแหน่งเพื่อวัดความสูงของ Cup และ Disc”

หน้าจอสำหรับการทดสอบความรู้ก่อนและหลังเรียน ในการทดสอบ นักศึกษาแพทย์ดูภาพแล้วกรอกข้อมูลตัวเลขในช่อง CDR ที่กำหนด และเปลี่ยนภาพโดยคลิกปุ่ม Next ดังภาพที่ 8 เมื่อทำการทดสอบเรียบร้อยแล้ว โปรแกรมจะแสดงผลการทดสอบ ดังภาพที่ 9 นอกจากนี้ ยังมีการบันทึกข้อมูล และสามารถคูณผลได้โดยผ่าน โปรแกรม Excel ดังภาพที่ 10



ภาพที่ 8: หน้าจอแสดงการใช้งานทดสอบก่อนและหลังเรียน

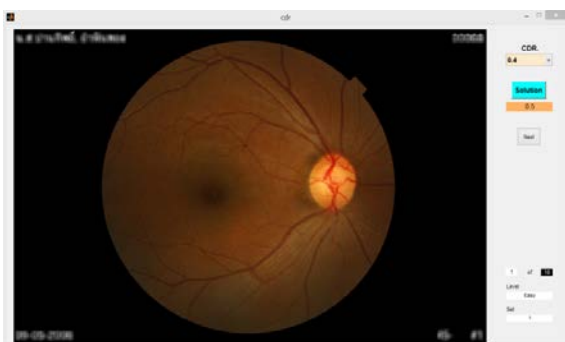


ภาพที่ 9: หน้าจอแสดงผลการทดสอบ

	A	B	C	D	E	F	G	H	I	J	K
1	test01	test02	test03	test04	test05	test06	test07	test08	test09	test10	username
2	0	0	0	0	1	0	1	0	1	0	5600000010

ภาพที่ 10: ผลการบันทึกข้อมูลผลการทดสอบผ่านโปรแกรม Excel

หน้าจอสําหรับการฝึกทักษะการประมาณค่าอัตราส่วนความสูงแนวตั้งระหว่าง Cup และ Disc ในหน้าจอนี้ นักศึกษาแพทย์สามารถฝึกหัดโดยการดูภาพแล้วตอบค่า CDR และดูคำตอบของอาจารย์แพทย์เพื่อเปรียบเทียบได้ทันที ดังภาพที่ 11



ภาพที่ 11: หน้าจอแสดงสําหรับการฝึกทักษะการประมาณค่า CDR

หน้าจอสําหรับฝึกทักษะการกำหนดตำแหน่งเพื่อวัดความสูงแนวตั้งของ Cup และ Disc หน้าจอนี้ใช้สําหรับฝึกกำหนดตำแหน่งเพื่อวัดส่วนสูงแนวตั้งของ Cup และ Disc โปรแกรมจะแสดงค่าความสูงในแนวตั้งของ Cup และ Disc ตามที่ User

กำหนด และผู้ใช้สามารถดูคำตอบเพื่อเปรียบเทียบกับคำตอบของอาจารย์แพทย์ได้ทันที ดังภาพที่ 12



ภาพที่ 12: หน้าจอแสดงสําหรับการฝึกทักษะการกำหนดตำแหน่งเพื่อวัดส่วนสูงแนวตั้งของ Cup และ Disc (CDR)

3.4 การทดสอบระบบ

ผู้วิจัยได้ทดสอบระบบโดยแบ่งการทดสอบออกเป็น 2 ขั้นตอน ดังนี้

1) การทดสอบขั้นแอลฟา (Alpha Testing) เป็นการทดสอบความสมบูรณ์ของโปรแกรมโดยผู้วิจัยเอง โดยการป้อนข้อมูลเข้าไปในโปรแกรมเพื่อให้ทราบว่โปรแกรมมีข้อผิดพลาดอะไรเกิดขึ้นบ้าง แล้วปรับปรุงโปรแกรมให้ดีขึ้นเพื่อนำไปทดสอบในขั้นต่อไป

2) การทดสอบขั้นเบต้า (Beta Testing) เป็นการทดสอบความสมบูรณ์ของโปรแกรมโดยใช้แบบสอบถามเพื่อการประเมินความพึงพอใจในการใช้งาน ซึ่งทำการประเมินกลุ่มตัวอย่างใช้แบบทดสอบความพึงพอใจ กลุ่มตัวอย่างเป็นนักศึกษาแพทย์เฉพาะทางด้านจักษุวิทยา คณะแพทยศาสตร์โรงพยาบาลรามาธิบดี ชั้นปีที่ 1-3 ชั้นปีละ 5 คน รวมทั้งสิ้น 15 คน

3.5 การประเมินคุณภาพระบบ

กลุ่มตัวอย่างประเมินระบบโดยใช้แบบสอบถามเพื่อประเมินความพึงพอใจในหัวข้อต่อไปนี้

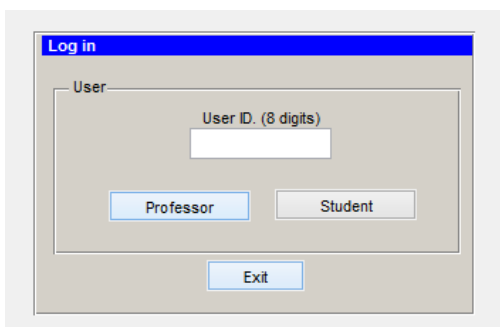
- 1) การออกแบบหน้าจอการใช้งาน -- เพื่อประเมินความเหมาะสมต่อการใช้งาน
- 2) ความหลากหลายของรูปภาพที่ใช้ทดสอบ -- เพื่อประเมินความครอบคลุมของเนื้อหา
- 3) ขั้นตอนการใช้งาน -- เพื่อประเมินการใช้งานมีความสะดวกเข้าใจง่าย
- 4) ประสิทธิภาพ -- เพื่อประเมินว่าระบบตรงต่อความต้องการใช้งาน

5) จำนวนใช้งานเฉลี่ย -- โดยนับจากจำนวนชุดการฝึกทักษะก่อนการทดสอบหลังเรียน

4. ผลการดำเนินการ

4.1 ผลการพัฒนาระบบ

ผลการวิจัยสามารถพัฒนาบทเรียนช่วยสอน เรื่อง การอ่านคำอัตราส่วนแนวตั้งระหว่าง Cup และ Disc (CDR) จากภาพถ่ายจอประสาท มีหน้าจอกำหนดการใช้งาน ดังภาพที่ 13



ภาพที่ 13: หน้าจอกำหนดการใช้งานบทเรียนช่วยสอน

4.2 ผลการประเมินผลสัมฤทธิ์ทางการศึกษา

ในการประเมินผลสัมฤทธิ์ 3 แบบคือ การประเมินผลสัมฤทธิ์ทางการเรียนรู้ก่อนและหลังเรียน ผลสัมฤทธิ์ของการประมาณค่า CDR และผลสัมฤทธิ์ของความแม่นยำในการกำหนดตำแหน่งเพื่อวัดความสูงของ Cup และ Disc บทเรียนแต่ละบทประกอบด้วยแบบทดสอบจำนวน 10 ภาพ ในการประเมินว่าถูกต้อง ค่า CDR สามารถคลาดเคลื่อนได้ (+/- 0.1) ได้ผลดังนี้

1) ผลสัมฤทธิ์ของการเรียนรู้ก่อนและหลังเรียน

ในการวิเคราะห์ผลสัมฤทธิ์ของการเรียนรู้ ผู้วิจัยได้นำบทเรียนช่วยสอนที่พัฒนานี้ไปใช้ประกอบการเรียนรู้กับกลุ่มตัวอย่าง วัดผลโดยใช้โปรแกรมทดสอบความรู้ก่อนและหลังการใช้บทเรียนคอมพิวเตอร์ช่วยสอนนี้ ทดสอบด้วยแบบทดสอบเดียวกันทั้งก่อนและหลังเรียน

ผู้วิจัยได้นำคะแนนทั้งก่อนและหลังเรียนของนักศึกษาแพทย์ ทั้ง 15 คน คำนวณด้วยสถิติ t-test ผลการคำนวณได้ดังตารางที่ 1

ตารางที่ 1: ผลการเปรียบเทียบผลสัมฤทธิ์ของการเรียนรู้

กลุ่มตัวอย่าง	N	\bar{X}	ค่า t
ก่อนเรียน	15	4.20	17.49
หลังเรียน	15	8.33	

* นัยสำคัญทางสถิติที่ระดับ .05

จากตารางที่ 1 ค่าเฉลี่ยของจำนวนภาพที่ผู้เรียนทำถูกต้องจากแบบทดสอบ 10 ภาพก่อนเรียนมีค่า 4.20 และค่าเฉลี่ยหลังเรียนมีค่า 8.33 มีค่าสถิติ t-test 17.49 จึงสรุปได้ว่าผลสัมฤทธิ์ทางการเรียนรู้หลังเรียนสูงกว่าก่อนเรียนอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และคะแนนเฉลี่ยหลังเรียนสูงกว่าก่อนเรียน

2) ผลสัมฤทธิ์ของการประมาณค่า CDR

ในการวัดผลสัมฤทธิ์ของการประมาณค่า CDR ผู้วิจัยนำผลของการเรียนในบทเรียนแรกและบทเรียนสุดท้ายของผู้เรียนแต่ละคนมาเปรียบเทียบกันดังผลในตารางที่ 2

ตารางที่ 2: ผลการเปรียบเทียบผลสัมฤทธิ์ของการประมาณค่า CDR

กลุ่มตัวอย่าง	N	\bar{X}	ค่า t
ก่อนเรียน	15	4.93	16.14
หลังเรียน	15	7.88	

จากตารางที่ 2 ค่าเฉลี่ยของจำนวนภาพที่ผู้เรียนทำถูกต้องจากแบบทดสอบ 10 ภาพก่อนเรียนมีค่า 4.93 และค่าเฉลี่ยหลังเรียนมีค่า 7.87 ค่าสถิติ t-test 16.14 จึงสรุปได้ว่าผลสัมฤทธิ์ของการประมาณค่า CDR หลังเรียนสูงกว่าก่อนเรียนอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และคะแนนเฉลี่ยหลังเรียนสูงกว่าก่อนเรียน

3) ผลสัมฤทธิ์ของความแม่นยำในการกำหนดตำแหน่งเพื่อวัดความสูงของ Cup และ Disc

เนื่องจากการกำหนดตำแหน่งที่ผู้เรียนกำหนดเพื่อวัดความสูงของ Cup และ Disc จำเป็นต้องวัดเป็นคู่ การวัดความแม่นยำในการกำหนดจุดจึงมีความยุ่งยาก ดังนั้น ในงานวิจัยนี้จึงประเมินความแม่นยำในการกำหนดตำแหน่งด้วยการประเมินค่า CDR ที่ได้จากการกำหนดตำแหน่งแทน ผลการประเมินที่วัดได้แสดงไว้ในตารางที่ 3

ตารางที่ 3: ผลการเปรียบเทียบผลสัมฤทธิ์ของความแม่นยำในการกำหนดตำแหน่งผ่านค่า CDR ที่คำนวณได้จากการกำหนดตำแหน่ง

กลุ่มตัวอย่าง	N	\bar{X}	ค่า t
ก่อนเรียน	15	5.07	22.10
หลังเรียน	15	8.27	

จากตารางที่ 3 ค่าเฉลี่ยของจำนวนภาพที่ผู้เรียนทำถูกต้องจากแบบทดสอบ 10 ภาพก่อนเรียนมีค่า 5.07 และค่าเฉลี่ยหลังเรียนมีค่า 8.27 ค่าสถิติ t-test 22.10 จึงสรุปได้ว่าผลสัมฤทธิ์ของ

ความแม่นยำในการกำหนดตำแหน่งและประมาณจากค่า CDR หลังเรียนสูงกว่าก่อนเรียนอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และคะแนนเฉลี่ยหลังเรียนสูงกว่าก่อนเรียน

4.3 ผลการประเมินความพึงพอใจของผู้ใช้งานระบบ

หลังจากนักศึกษาแพทย์กลุ่มตัวอย่างเรียนรู้โดยใช้บทเรียนช่วยสอนแล้ว ผู้วิจัยได้ประเมินความพึงพอใจของผู้วิจัยได้ประเมินความพึงพอใจโดยให้คะแนนระดับความพึงพอใจ น้อยถึงมาก (1 – 5 คะแนน) โดยผลการประเมิน ผู้ใช้งานมีความพึงพอใจการใช้งานโปรแกรมโดยรวมอยู่ในระดับดีมาก (\bar{X} = 4.55, S.D. = 0.61) และได้ผลการประเมินแสดงดังตารางที่ 4

ตารางที่ 4 ผลการประเมินความพึงพอใจ

รายการ	\bar{X}	S.D.
1.การออกแบบหน้าจอ	4.60	0.63
2.ความหลากหลายของรูปภาพที่ใช้ทดสอบ	4.53	0.52
3.ขั้นตอนการใช้งานเข้าใจง่าย	4.53	0.64
4.ความตรงต่อความต้องการใช้งาน.	4.60	0.51
5.จำนวนการใช้งานเฉลี่ย	4.47	0.74
เฉลี่ยรวม	4.55	0.61

จากตารางที่ 4 ผลการประเมินความพึงพอใจของนักศึกษาแพทย์ เมื่อพิจารณารายด้าน การออกแบบหน้าจอ นักศึกษาแพทย์ส่วนใหญ่มีความพึงพอใจอยู่ในระดับดีมาก (\bar{X} = 4.60, S.D. = 0.63) ด้านความหลากหลายของรูปภาพที่ใช้ทดสอบส่วนใหญ่มีความพึงพอใจอยู่ในระดับดีมาก (\bar{X} = 4.53, S.D. = 0.52) ด้านความง่ายของขั้นตอนการใช้งาน ส่วนใหญ่มีความพึงพอใจอยู่ในระดับดีมาก (\bar{X} = 4.53, S.D. = 0.64) ด้านความตรงต่อการใช้งาน ส่วนใหญ่มีความพึงพอใจอยู่ในระดับดีมาก (\bar{X} = 4.60, S.D. = 0.51) และมีจำนวนการใช้งานเฉลี่ย 4.47 ครั้ง

5. บทสรุปและข้อเสนอแนะ

การพัฒนาบทเรียนช่วยสอนมีการใช้เทคนิคการใช้ภาพเป็นตัวประสานกับผู้ใช้ Graphical User Interface (GUI) ซึ่งผลการพัฒนาระบบทำให้ผลสัมฤทธิ์ทางการเรียนหลังใช้บทเรียนช่วยสอนของนักศึกษาแพทย์ดีขึ้น และสามารถนำไปประยุกต์ใช้ได้จริง

อย่างไรก็ดี ผู้วิจัยเห็นว่าควรมีการพัฒนาโปรแกรมวัดค่าอัตราส่วนแนวตั้งระหว่าง Disc และ Cup จากภาพถ่ายจอประสาทตาแบบอัตโนมัติเพราะจะช่วยลดงานให้แพทย์และเพิ่ม

ความแม่นยำในการวินิจฉัยโรคต้อหินซึ่งทำให้ทั้งแพทย์และผู้ป่วยได้รับประโยชน์ร่วมกัน

เอกสารอ้างอิง

- [1] ยืน ภู่วรรณ และสมชาย นำประเสริฐชัย, “ไอซีทีเพื่อการศึกษาไทย” ซีเอ็ดดูเคชั่น 2546.
- [2] ถนอมพร เลหาจรัสแสง, “การออกแบบบทเรียนคอมพิวเตอร์ช่วยสอน” เอกสารประกอบการสอน ภาควิชาโสตทัศนศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- [3] How to Treat Glaucoma (2556) (ออนไลน์). สืบค้นจาก : <http://www.howtoimproveyeyesightwithouturgery.com/how-to-treat-glaucoma/> [1 กุมภาพันธ์ 2557]
- [4] ต้อหิน (2557) (ออนไลน์). สืบค้นจาก : <http://www.thaiglaucoma.org/article/ต้อหิน-glaucoma> [20 มกราคม 2557]
- [5] Lillian C. Lee, M.D.(2556). WHAT IS GLAUCOMA (ออนไลน์). สืบค้นจาก : <http://www.lcleemd.com/glaucoma.html> [15 ธันวาคม 2556]
- [6] ทิพย์มณฑา สดชื่น, “การพัฒนาหนังสืออิเล็กทรอนิกส์ เรื่องการถ่ายภาพเบื้องต้น” วิทยานิพนธ์ ศศ.ม. สงขลา : มหาวิทยาลัยสงขลานครินทร์ 2544.
- [7] อินทรา ชูศรีทอง, “การพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนระบบมัลติมีเดีย วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6” รายงานการศึกษาค้นคว้าอิสระ กศ.ม. มหาวิทยาลัยมหาสารคาม 2543.

ปัจจัยที่ส่งผลต่อเด็กออทิสติกในมุมมองของผู้ปกครอง และการใช้ชุดฝึกมัลติมีเดียเกม ศิลปะ สร้างกระบวนการเรียนรู้เรื่องสี โดยแบ่งกลุ่มอาการเด็กออทิสติก

Factors Affecting Autistic Children in Parents' Perspectives and the Use of Multimedia Art Game Set to Create Color Learning Process Divided by Groups of Symptoms

กนิศร จีระโทก (Kanisorn Jeekratok)

สาขาวิชาวิทยาการคอมพิวเตอร์ และเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏอุดรธานี

Kanisorn_je@hotmail.com

บทคัดย่อ

การวิจัยนี้มีจุดมุ่งหมายเพื่อศึกษามุมมองของผู้ปกครองต่อปัจจัยที่มีผลต่อเด็กออทิสติกที่เข้ารับบริการที่ศูนย์การศึกษาพิเศษประจำจังหวัดหนองคาย เครื่องมือที่ใช้คือ แบบสังเกตและรวบรวมข้อมูลโดยการสัมภาษณ์กลุ่มผู้ปกครองเป้าหมาย การเฝ้าดูพฤติกรรมเด็กออทิสติก และแบ่งกลุ่มอาการของเด็กก่อน และหลังการได้รับการฝึกโดยชุดฝึกมัลติมีเดียเกม ศิลปะของเด็กที่ศูนย์การศึกษาพิเศษประจำจังหวัดหนองคาย วิเคราะห์โดยใช้สถิติพื้นฐาน หาค่าร้อยละ ค่า *t-test* และการวิเคราะห์เนื้อหา จากการสัมภาษณ์ผู้ปกครองกลุ่มตัวอย่างแบบเจาะจงจำนวน 12 คนในศูนย์การศึกษาพิเศษประจำจังหวัดหนองคาย พบว่า ปัจจัยที่มีผลต่อการสร้างพัฒนาการและการเรียนรู้ของเด็กในความคิดเห็นของบิดามารดา ได้แก่ ความรู้เกี่ยวกับโรคออทิสติก อายุ ระดับการศึกษา รายได้ เวลาที่ได้รับการฝึก และการปรับตัวเพื่อให้ยอมรับความจริง โดยส่วนใหญ่มารดาจะสามารถปรับตัวและยอมรับได้เร็วกว่า บิดา และพร้อมที่จะเปิดรับวิธีการในการรักษามากกว่า และคาดหวังว่าลูกจะสามารถปรับตัวเข้ากับสังคม หรือเรียนในโรงเรียนร่วมได้ในอนาคต จากการเฝ้าสังเกตกลุ่มตัวอย่างแบบเจาะจงจำนวน 15 คน ที่ได้รับการอนุญาตจากผู้ปกครองเด็กเพื่อให้สามารถร่วมในการวิจัย ใช้ระยะเวลา 7 สัปดาห์ ในการเฝ้าสังเกตก่อน และหลังการใช้ชุดฝึกมัลติมีเดียเกม ศิลปะ โดยมี เด็กผู้ชาย 8 คน เด็กผู้หญิง 7 คน อายุระหว่าง 8-14 ปี ที่เข้ารับบริการในศูนย์การศึกษาพิเศษประจำจังหวัดหนองคาย ผลการสังเกตปรากฏว่า ผู้วิจัยแบ่งกลุ่มอาการออกเป็น 3 กลุ่มอาการคือ 1.กลุ่มที่มีอาการและพฤติกรรมรุนแรงมาก คิดเป็นร้อยละ 33.33 % 2.กลุ่มที่มีอาการและพฤติกรรมรุนแรงปานกลาง คิด

เป็นร้อยละ 46.66 % 3.กลุ่มที่มีอาการและพฤติกรรมรุนแรงน้อย คิดเป็นร้อยละ 20 % โดยในกลุ่มอาการที่ 1 นั้น ไม่สามารถใช้ชุดฝึกมัลติมีเดียเกม ศิลปะ ในการสร้างกระบวนการเรียนรู้ได้ แต่ยังสามารถใช้แผนกิจกรรมศิลปะ ในการพัฒนากระบวนการเรียนรู้เรื่องสีในเด็กได้ กลุ่มที่ 2 และกลุ่มที่ 3 มีจำนวนเด็ก 10 คน โดยมี เด็กผู้ชาย 6 คน เด็กผู้หญิง 4 คน จากการเฝ้าสังเกต พบว่า เด็กสามารถบอกชื่อสีได้ถูกต้องมากขึ้น ความสามารถในการระบายสีโดยให้อยู่ในกรอบที่กำหนด เด็กสามารถบังคับมือ (กล้ามเนื้อมัดเล็ก) ทำให้ดีโดยระบายสีออกนอกกรอบน้อยลง และความสามารถในการแยกความแตกต่างระหว่างสีเข้มและสีอ่อนเด็กมีพัฒนาการที่ดีขึ้น เมื่อมองแยกกลุ่มตามกลุ่มอาการของเด็ก พบว่า กลุ่มที่มีอาการและพฤติกรรมรุนแรงปานกลางและรุนแรงน้อย นั้นสามารถใช้ชุดฝึกมัลติมีเดียเกม ศิลปะในการช่วยสร้างกระบวนการในการพัฒนาความรู้ได้

คำสำคัญ: ปัจจัย มุมมองผู้ปกครอง เด็กออทิสติก ชุดฝึกมัลติมีเดียเกม ศิลปะ

Abstract

The objective of this research is to study the perspectives of parents towards the factors which affect autistic children who use the service at the Special Education Nongkhai Center. The equipment used is the observation form, and the data compilation is done by interviewing the target group of parents, observing autistic children's behavior and dividing groups of symptoms before and after the training by multimedia art game set of children at the Special Education Nongkhai Center. The analysis is done by basic statistics,

percentage, *t*-test and content analysis. According to the focused interview with 12 parents in the sample group at the Special Education Nongkhai Center, it was discovered that the factors which affected the children's development and learning process in parents' perspectives were the knowledge about autism, age, educational level, income, period of training and adaptation to accept the truth. Mostly, mothers were able to adapt themselves and accept the truth faster than fathers. Also, they were more ready to open up to the methods of treatment and expect that their children would be able to adapt themselves to live in the society or study in normal schools in the future. According to the focused observation of the sample group of 15 children who were allowed by their parents to participate in the research which took 7 weeks to observe 8 boys and 7 girls aged between 8-14 years old from the Special Education Nongkhai Center before and after the use of multimedia art game set, the researcher was able to divide the symptoms into 3 groups, which were 1. the group with severe symptoms and behavior which was considered 33.33 %, 2. the group with moderate symptoms and behavior which was considered 46.66 % and 3. the group with mild symptoms and behavior which was considered 20 %. Children who belonged to Group 1 could not use the multimedia art game set to build up their learning process. Still, they could follow the art activity plan to develop their color learning process in children. In Group 2 and Group 3, 10 children (6 boys and 4 girls) could specify the colors more accurately. As for the ability to color pictures inside the frame, it was found out that the children could control their hands and fingers (fine motor development) quite well as they colored outside the frame less. Also, their ability to differentiate dark and light colors improved. When considering children according to their groups of symptoms, it was discovered that the children whose symptoms and behavior were moderate and mild were able to use the multimedia art game set to help create their learning process.

Keywords: factors, parents' perspectives, autistic children, multimedia art game set

1. บทนำ

โรคออทิสติก(Autistic Disorder) หรือออทิสซึม(Autism) เป็นความบกพร่องของพัฒนาการรูปแบบหนึ่งที่มีลักษณะเฉพาะ นับเป็นเวลานาน ที่นักวิชาการรู้จักโรคนี้ และพยายามศึกษา รายละเอียดต่างๆ ของโรค แต่ยังไม่สามารถหาสาเหตุของโรค ได้ชัดเจน[1] ออทิสติกพบได้ในเด็กทั่วโลก และพบในเด็กชาย มากกว่าเด็กหญิงถึง 4 เท่าจากการวินิจฉัยและสถิติของสมาคมจิตแพทย์อเมริกันครั้งที่ 3 และ 4 พบว่ามีเด็กออทิสติก 4-5 ราย ในประชากรเด็ก 10,000 ราย[2] ในสหรัฐอเมริกามีคนประมาณ 500,000-1,500,000 คน ที่มีอาการผิดปกติแบบออทิสติก ซึ่งนับได้ว่า โรคออทิสติก เป็นหนึ่งในความผิดปกติทางพัฒนาการที่ พบบ่อยที่สุดในปัจจุบัน[3] ในประเทศไทย พบว่า สถิติของ ประชากรจำนวน 10,000 คน จะพบบุคคลออทิสติก 4-5 คน และพบว่า ในเด็กอายุต่ำกว่า 1 ปี จำนวน 10,000 คนจะพบถึง 21 คน ซึ่งนับได้ว่ามีความชุกถึง 36 คนใน 10,000 คน[4] การดูแลเด็กออทิสติก จึงจำเป็นต้องใช้ความเข้าใจ ความอดทน รวมทั้งต้องมีความรู้เกี่ยวกับการปฏิบัติ การดูแลและให้ความช่วยเหลือเด็กได้อย่างถูกต้อง[5] ความพร้อมในการดูแลจึงเป็น ปัจจัยที่จำเป็น ไม่ว่าจะเป็นทางด้านร่างกาย และจิตใจ โดยเฉพาะอย่างยิ่งความพร้อมในด้านจิตใจ[6] การศึกษาวิจัย ครั้งนี้ผู้วิจัยใช้การบูรณาการชุดฝึกมัดติมีเดียเกม ศิลปะ และแผนการทำกิจกรรมศิลปะ โดยชุดฝึกมัดติมีเดียเกม เป็นกิจกรรมเสริมขึ้นเพื่อพัฒนาสมอง และเพื่อตั้งให้เด็กเกิดความสนใจ[7] กิจกรรมศิลปะนั้นจะช่วยทำให้เด็กมีสมาธิ เพราะต้องจดจ่ออยู่กับกิจกรรมที่เด็กกำลังทำ[8] ใช้แบบสัมภาษณ์ ผู้ปกครอง สังเกตพฤติกรรมเด็กก่อน-หลัง การใช้ชุดฝึกมัดติมีเดียเกม ศิลปะ และแผนการทำกิจกรรมศิลปะ แบบทดสอบการบอกชื่อสี การระบายสี การแยกสี ในการเก็บรวบรวมข้อมูลเพื่อให้ได้ข้อมูลที่เป็นแนวทาง และวิธีเพื่อที่จะนำไปพัฒนา และช่วยเหลือเด็กออทิสติกต่อไป

2. วัตถุประสงค์ของการวิจัย

2.1 เพื่อศึกษาปัจจัยที่มีผลต่อเด็กออทิสติกในมุมมองของผู้ปกครองที่เข้ารับบริการที่ศูนย์การศึกษาพิเศษประจำจังหวัดหนองคาย

2.2 เพื่อศึกษาความสามารถการจำแนกสีต่าง ๆ ในชีวิตประจำวันโดยใช้ชุดฝึกมัดติมิเดียเกม ศิลปะ โดยแบ่งกลุ่มอาการของเด็ก

3. เครื่องมือที่ใช้ในงานวิจัย

แบบสัมภาษณ์ผู้ปกครอง, ชุดฝึกมัดติมิเดียเกม, กิจกรรมศิลปะ, แบบทดสอบการบอกชื่อสี การระบายสี การแยกสี และการสังเกตพฤติกรรมและประเมินพฤติกรรมเด็กก่อนและหลัง โดยทำการตรวจสอบแบบสามเส้า (triangulation) และบันทึกผลพฤติกรรมเด็กโดยผู้วิจัย ครูฝึก และผู้ปกครองเด็ก

4. การเก็บรวบรวมข้อมูล

กลุ่มตัวอย่าง ได้แก่ ผู้ปกครองเด็กออทิสติก แบบเจาะจงจำนวน 12 คน และเฝ้าพฤติกรรมเด็กออทิสติกร่วมกับครูฝึกศูนย์การศึกษาพิเศษประจำจังหวัดหนองคาย จำนวนเด็ก 15 คน โดยเป็นเด็กผู้ชาย 8 คน และเด็กผู้หญิง 7 คน มีอายุระหว่าง 8-14 ปี ที่ผู้ปกครองนำมาบริการบำบัดที่ศูนย์การศึกษาพิเศษประจำจังหวัดหนองคาย และได้รับการอนุญาตจากผู้ปกครองเด็ก เพื่อให้สามารถร่วมในการวิจัย โดยใช้แบบสอบถาม แบบสังเกตพฤติกรรมเด็กออทิสติก โดยให้เด็กใช้ชุดฝึกมัดติมิเดียเกม ศิลปะ สัปดาห์ละ 3 ครั้ง ครั้งละ 1 ชั่วโมง รวม 7 สัปดาห์

5. ผลการวิจัย

การวิจัยครั้งนี้ได้แบ่งการแสดงผลออกเป็น 4 ส่วน ดังนี้

5.1 การสัมภาษณ์ผู้ปกครองกลุ่มตัวอย่างแบบเจาะจงจำนวน 12 คน เกี่ยวกับปัจจัยที่มีผลต่อการสร้างพัฒนาและการเรียนรู้ในเด็กออทิสติก

ศูนย์การศึกษาพิเศษประจำจังหวัดหนองคายได้ใช้รูปแบบการสอนตามโปรแกรม TEACCH (Treatment and Education of Autistic and related Communication handicapped CHildren) พัฒนาโดย Dr. Eric Schopler เน้นการสอนอย่างมีระบบ ขั้นตอน และการจัดสภาพแวดล้อมให้เหมาะสมกับเด็กเป็นหัวใจสำคัญ โดยมีการจัดห้องเรียนให้เป็นระบบ จัดของเป็นหมวดหมู่ จัดตารางเวลากิจกรรมต่างๆ แน่นนอน และการพัฒนาการเรียนรู้ แบบมอนเตสซอรี(Montessori)[9] พัฒนาโดย

Dr. Maria Montessori คือ หลักการที่คำนึงถึงเด็ก ความต้องการของเด็กในการเรียน ได้มีการเตรียมสิ่งแวดล้อมให้เด็กได้ทำงานด้วยตนเอง สิ่งแวดล้อม และเครื่องมืออุปกรณ์ต่างๆ นั้นได้จัดระบบไว้เพื่อให้เด็กได้พัฒนาตนเอง โปรแกรมจัดเอาไว้ให้เด็กได้เป็นผู้เรียนที่มีอิสระ การควบคุมความผิดพลาดในการทำงานก็ด้วยการใช้วัสดุ เหล่านั้นเอง และสิ่งแวดล้อมที่จัดเอาไว้ให้ตัวเองเป็นตัวที่ทำให้เด็กมีอิสระ ทางศูนย์การศึกษาพิเศษประจำจังหวัดหนองคายใช้รูปแบบการสอน “โดยใช้ประสาทสัมผัสและการเคลื่อนไหว คือ การเรียนปนเล่น” วิธีการสอนเริ่มจากการสังเกตเด็ก ศึกษาพัฒนาการของเด็ก ความต้องการและความสนใจของเด็ก จากการสัมภาษณ์ผู้ปกครองเด็กพบว่า ปัจจัยที่มีผลต่อการสร้างพัฒนาการและการเรียนรู้ของเด็กในบิดามารดา กลุ่มตัวอย่างแบบเจาะจงจำนวน 12 คน ได้แก่ ความรู้เกี่ยวกับโรคออทิสติก อายุ ระดับการศึกษา รายได้ เวลาที่ได้รับบริการฝึก และการปรับตัวเพื่อให้ยอมรับความจริง โดยส่วนใหญ่มารดาจะสามารถปรับตัวและยอมรับได้เร็วกว่า บิดา และพร้อมที่จะเปิดรับวิธีการในการรักษามากกว่า และคาดหวังว่าลูกจะสามารถปรับตัวเข้ากับสังคม หรือเรียนในโรงเรียนร่วมได้ในอนาคต

5.2 การเฝ้าสังเกตพฤติกรรมและอาการของเด็กออทิสติกกลุ่มตัวอย่างแบบเจาะจง จำนวน 15 คน โดยการสังเกตพฤติกรรมและอาการของเด็ก ร่วมกับครูฝึกศูนย์การศึกษาพิเศษประจำจังหวัดหนองคาย

สามารถจัดแบ่งกลุ่มอาการออกเป็น 3 กลุ่ม ดังนี้

1.กลุ่มที่มีอาการและพฤติกรรมรุนแรงมาก จำนวน 5 คน คิดเป็นร้อยละ (33.33%) เด็กกลุ่มนี้ไม่สามารถแยกสีต่างๆ ได้ ส่วนพฤติกรรมทางสังคม เช่นการเก็บของใช้ การช่วยเหลือแบ่งปัน ไม่สามารถทำได้เองต้องมีครูผู้ฝึกคอยช่วยเหลือตัวอย่างเช่น เมื่ออยู่ในห้องเรียนจะมีการเรียนโดยใช้อุปกรณ์ต่างๆ เพื่อดึงดูดความสนใจของเด็ก เมื่อครูผู้ฝึกให้อุปกรณ์ในการเรียนกับเด็กกลุ่มนี้ เขาก็จะไม่ยอมให้เด็กคนอื่นเข้ามาใกล้กับอุปกรณ์ของตนเอง หากเข้ามาใกล้ก็จะหยิบอุปกรณ์แล้วเดินหนี หากยังเดินตามบางครั้งเขาก็จะโวยวาย และอาจทำร้ายเพื่อนที่เดินเข้ามาใกล้ได้ และพฤติกรรมการปฏิบัติตามกฎระเบียบ เช่น การร่วมกลุ่ม รอการใช้สิ่งของในห้องกิจกรรม เด็กกลุ่มนี้ไม่สามารถทำได้ ตัวอย่างเช่น เมื่อมีการเข้าไปในห้องกิจกรรม

หากเด็กในกลุ่มนี้ต้องการเล่นของเล่น เด็กก็จะตรงไปที่ของเล่น โดยไม่สนใจว่าจะมีเด็กคนไหนเล่นอยู่ จะตรงเข้าไปแย่งของเล่นมาเล่น หากไม่ได้ ก็จะโวยวายเสียงดัง หรือในบางครั้งก็จะเข้าไปทำร้ายเด็กที่กำลังเล่นอยู่เพื่อให้ได้ของเล่นที่ตนเองต้องการ ไม่สามารถใช้ชุดฝึกมัดติมิเดียเกม ศิลปะ ในการสร้างกระบวนการเรียนรู้ได้ แต่ยังสามารถใช้แผนกิจกรรมศิลปะในการพัฒนากระบวนการเรียนรู้เรื่องสีในเด็กได้

2.กลุ่มที่มีอาการและพฤติกรรมรุนแรงปานกลาง จำนวน 7 คน คิดเป็นร้อยละ (46.66%) เด็กกลุ่มนี้มีความสามารถในการแยกสีได้เป็นบางสี ตัวอย่างเช่น สีเขียว สีฟ้า เป็นต้น พฤติกรรมทางสังคม เช่นการเก็บของใช้ การช่วยเหลือแบ่งปัน สามารถทำได้บ้าง และพร้อมที่จะรับความรู้จากครูผู้ฝึกสอนได้ โดยใช้การสอนแบบซ้ำๆ ตัวอย่างเช่น เด็กในกลุ่มนี้อยู่ในห้องเรียน ครูผู้ฝึกสอนให้ดูรูปภาพ หรือตุ๊กตารูปสัตว์ เมื่อมีเพื่อนเข้ามาดูด้วย เด็กกลุ่มนี้จะไม่โวยวาย และหิบบของเดินหนี แต่จะนั่งอยู่เฉยๆ ให้เพื่อนได้ดูกับตนเอง ครูผู้ฝึกสอนบอกให้หิบบรูปภาพให้เพื่อนดู เด็กกลุ่มนี้สามารถหิบบรูปภาพได้ตรงตามที่ครูบอก และส่งไปให้เพื่อนดูได้ และพฤติกรรมกรปฏิบัติตามกฎระเบียบ เช่น การร่วมกลุ่ม รอคารใช้สิ่งของในห้องกิจกรรม เด็กกลุ่มนี้สามารถเล่นเป็นกลุ่มได้โดยต้องมีครูผู้ฝึกสอนดูแล มีการโวยวายในบางครั้งที่อารมณ์เด็กอาจจะไม่ดี ตัวอย่างเช่น เมื่อเด็กได้เข้าไปที่ห้องกิจกรรมเมื่อเค้าอยากเล่นของเล่นอะไร เด็กจะคิดว่ามีใครเล่นของเล่นชิ้นนั้นอยู่ก่อนหรือเปล่า หากมีเค้าจะรอนกว่าเพื่อนจะเล่นเสร็จแล้วจึงเดินไปหิบบของเล่นมาเล่น แม้จะมีการโวยวาย หรือเดินออกจากห้องฝึกบางเป็นบางครั้ง

3.กลุ่มที่มีอาการและพฤติกรรมรุนแรงน้อย จำนวน 3 คน คิดเป็นร้อยละ (20%) เด็กกลุ่มนี้มีความสามารถจดจำสีที่ครูผู้ฝึกสอน ทำการสอนได้ในระดับดี เช่น สีชมพู, สีแดง, สีเหลือง, สีน้ำเงิน, สีเขียว เป็นต้น และเด็กบางคนในกลุ่มนี้มีความสามารถทางภาษาอังกฤษอยู่ในระดับดี รู้คำศัพท์ภาษาอังกฤษและสามารถพูดตอบโต้เป็นประโยคได้ พฤติกรรมทางสังคม เช่นการเก็บของใช้ การช่วยเหลือแบ่งปัน สามารถทำได้โดยที่ครูผู้ฝึกแนะนำให้ทำเพียงครั้งเดียว เด็กจะจำและรู้ว่าต้องทำอะไรในครั้งต่อไป เวลาอยู่ในห้องเรียนเด็กสามารถช่วยครูผู้ฝึกในการหิบบอุปกรณ์ต่าง ๆ มาเองได้ และนำไปเก็บได้ถูกต้องเมื่อใช้เสร็จเรียบร้อย และพฤติกรรมกรปฏิบัติตาม

กฎระเบียบ เช่น การร่วมกลุ่ม รอคารใช้สิ่งของในห้องกิจกรรม เด็กกลุ่มนี้สามารถเล่นเป็นกลุ่มได้ รู้จักการรอคอย และเด็กบางคนในกลุ่มนี้จะชอบการเข้ากลุ่มเล่นกับคนอื่น แม้จะมีการโวยวาย หรือเดินออกจากห้องฝึกบางเป็นบางครั้ง

5.3 ความสามารถในการจำแนกสีต่าง ๆ ในชีวิตประจำวัน แบ่งกลุ่มอาการเด็ก โดยใช้ชุดฝึกมัดติมิเดียเกมศิลปะ และแผนกิจกรรมศิลปะ

โดยแบ่งเป็น 3 ด้าน ได้แก่ 1.การบอกชื่อสี 2.การระบายสี 3. การแยกสี โดยแสดงเป็นคำร้อยละ ก่อน-หลังการใช้ชุดฝึกมัดติมิเดียเกม กิจกรรมศิลปะ และค่า t-test ผลการศึกษาความสามารถ โดยแยกกลุ่มอาการเด็กในกลุ่มที่ 2 กลุ่มที่มีอาการและพฤติกรรมรุนแรงปานกลาง และกลุ่มที่ 3 กลุ่มที่มีอาการและพฤติกรรมรุนแรงน้อย แสดงรายละเอียดดังในตารางที่ 1 – 3

ตารางที่ 1: เรื่องการบอกชื่อสี โทนสีแตกต่างกัน และโทนสีใกล้เคียงกัน แสดงเป็นคำร้อยละ และค่า t-test ดังนี้

เด็กคนที่	เพศ	อายุ	กลุ่มอาการ	โทนสีแตกต่างกัน			โทนสีใกล้เคียงกัน			ภาพรวมพัฒนาการ
				% ก่อนการฝึก	% หลังการฝึก	% ที่พัฒนาขึ้น	% ก่อนการฝึก	% หลังการฝึก	% ที่พัฒนาขึ้น	
A	หญิง	8	ปานกลาง	41	66	25	25	50	25	พอใช้
B	ชาย	9	ปานกลาง	25	41	16	33	50	16	ต้องปรับปรุง
C	ชาย	10	ปานกลาง	41	66	25	33	58	25	พอใช้
D	ชาย	12	ปานกลาง	25	41	16	16	41	25	ต้องปรับปรุง
E	ชาย	13	ปานกลาง	25	41	16	33	66	33	พอใช้
F	หญิง	14	ปานกลาง	33	66	33	50	75	50	ดีขึ้น
G	ชาย	14	ปานกลาง	25	41	16	25	50	25	ต้องปรับปรุง
H	หญิง	9	น้อย	25	75	50	33	83	50	ดีขึ้น
I	ชาย	10	น้อย	33	58	27	33	58	25	พอใช้
J	หญิง	11	น้อย	33	91	58	33	91	58	ดีขึ้น
t-test				5.83**			7.44**			

** P < .01

จากตารางที่ 1 ภาพรวมของการพัฒนาการในเรื่องการบอกชื่อสี โทนสีแตกต่างกันคิดเป็น 16% - 58% และโทนสีใกล้เคียงกันคิดเป็น 16% - 58% โดยเด็กที่มีภาพรวมพัฒนาต่ำสุด ได้แก่ เด็กผู้ชายอายุ 9 ปี มีภาพรวมพัฒนาการทั้ง 2 ด้านต่ำสุดที่ 18 % และ 18 % ตามลำดับ และเด็กที่มีภาพรวมพัฒนาสูงสุด ได้แก่ เด็กผู้หญิงอายุ 11 ปี ภาพรวมพัฒนาการทั้ง 2 ด้านสูงสุดที่ 58 % และ 58 % ตามลำดับ

ตารางที่ 2: เรื่องการระบายโทนสีแตกต่างกัน และโทนสีใกล้เคียงกัน แสดงเป็นคำร้อยละ ดังนี้

เด็กคนที่	เพศ	อายุ	กลุ่มอาการ	โทนสีแตกต่างกัน			โทนสีใกล้เคียงกัน			ภาพรวมพัฒนาการ
				% ก่อนการฝึก	% หลังการฝึก	% ที่พัฒนาขึ้น	% ก่อนการฝึก	% หลังการฝึก	% ที่พัฒนาขึ้น	
A	หญิง	8	ปานกลาง	36	54	18	27	45	18	ต้องปรับปรุง
B	ชาย	9	ปานกลาง	27	45	18	27	54	18	ต้องปรับปรุง
C	ชาย	10	ปานกลาง	36	90	54	36	72	36	ดีขึ้น
D	ชาย	12	ปานกลาง	27	45	18	27	54	27	พอใช้
E	ชาย	13	ปานกลาง	27	90	63	36	81	45	ดีขึ้น
F	หญิง	14	ปานกลาง	27	54	25	27	54	25	พอใช้
G	ชาย	14	ปานกลาง	18	36	18	27	45	18	ต้องปรับปรุง
H	หญิง	9	น้อย	27	54	27	36	63	27	พอใช้
I	ชาย	10	น้อย	27	45	18	36	54	18	ต้องปรับปรุง
J	หญิง	11	น้อย	36	63	27	36	54	18	พอใช้
t-test				5.59**			9.18**			

** P < .01

จากตารางที่ 2 ภาพรวมของการพัฒนาการในเรื่องการระบายสีโทนสีแตกต่างกันคิดเป็น 18% - 63% และโทนสีใกล้เคียงกันคิดเป็น 18% - 45% โดยเด็กที่มีภาพรวมพัฒนาต่ำสุด 3 คน ได้แก่ เด็กผู้ชายอายุ 14, 9 ปี และเด็กผู้หญิงอายุ 8 ปี มีภาพรวมพัฒนาการทั้ง 2 ด้านต่ำสุดที่ 18 % และ 18 % ตามลำดับ ทั้ง 3 คน และเด็กที่มีภาพรวมพัฒนาสูงสุด ได้แก่ เด็กผู้ชายอายุ 13 ปี ภาพรวมพัฒนาการทั้ง 2 ด้านสูงสุดที่ 63% และ 45 % ตามลำดับ

ตารางที่ 3: เรื่องการแยกสีโทนสีแตกต่างกันและโทนสีใกล้เคียงกัน แสดงเป็นคำร้อยละ ดังนี้

เด็ก คนที่	เพศ	อายุ	กลุ่มอาการ	โทนสีแตกต่างกัน		โทนสีใกล้เคียงกัน		ภาพรวม พัฒนาการ		
				% ก่อนการฝึก	% หลังการฝึก	% ที่พัฒนาขึ้น	% หลังการฝึก			
A	ญ	8	ปานกลาง	41	66	25	33	50	16	ดีเยี่ยม/สูง
B	ช	9	ปานกลาง	25	58	33	33	66	33	พอใช้
C	ช	10	ปานกลาง	41	75	33	33	66	33	พอใช้
D	ช	12	ปานกลาง	33	66	33	25	58	33	พอใช้
E	ช	13	ปานกลาง	33	50	16	33	66	33	ดีเยี่ยม/สูง
F	ญ	14	ปานกลาง	33	83	50	33	50	16	พอใช้
G	ช	14	ปานกลาง	25	41	16	25	41	16	ดีเยี่ยม/สูง
H	ญ	9	น้อย	33	83	50	33	66	33	ดีเยี่ยม
I	ช	10	น้อย	41	75	33	33	83	30	ดีเยี่ยม
J	ญ	11	น้อย	33	83	50	41	75	33	ดีเยี่ยม
test				8.41**			8.50**			

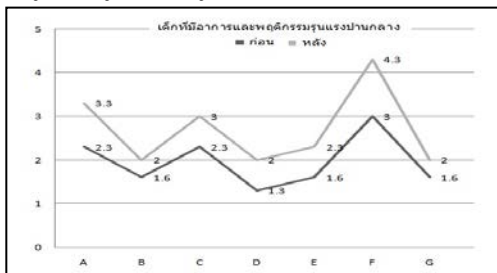
** P < .01

จากตารางที่ 3 ภาพรวมของการพัฒนาการในเรื่องการแยกสีโทนสีแตกต่างกันคิดเป็น 16% - 50% และการแยกสีโทนสีใกล้เคียงกันคิดเป็น 16% - 50% โดยเด็กที่มีภาพรวมพัฒนาต่ำสุด ได้แก่ เด็กผู้ชายอายุ 14 ปี มีพัฒนาการทั้ง 2 ด้านต่ำสุดที่ 16 % และ 16 % ตามลำดับ และเด็กที่มีภาพรวมพัฒนาสูงสุด 3 คน ได้แก่ เด็กผู้หญิงอายุ 9 และ 11 ปี มีพัฒนาการทั้ง 2 ด้านสูงสุด 50 % และ 33% ตามลำดับ เด็กผู้ชายอายุ 10 ปี พัฒนาการทั้ง 2 ด้านสูงสุด 33% และ 33% ตามลำดับ โดยทั้งหมดเป็นเด็กที่มีอาการและพฤติกรรมรุนแรงน้อย

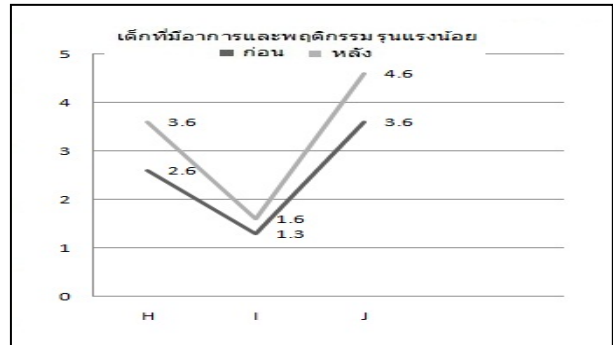
5.4 ความสามารถในการเรียนรู้เรื่องสี โดยแยกตามกลุ่มอาการ

ใช้การตรวจสอบแบบสามเส้า (triangulation) หาค่าเฉลี่ยก่อนและหลัง จากมุมมองผู้วิจัย ครูฝึก และผู้ปกครองเด็ก แสดงรายละเอียดดังในภาพที่ 1 – 6

ภาพที่ 1 และภาพที่ 2 แสดงเรื่องการบอกชื่อสี โทนสีแตกต่างกัน และโทนสีใกล้เคียงกัน แสดงค่าเฉลี่ยก่อนและหลัง จากมุมมองผู้วิจัย ครูฝึก และผู้ปกครองเด็ก

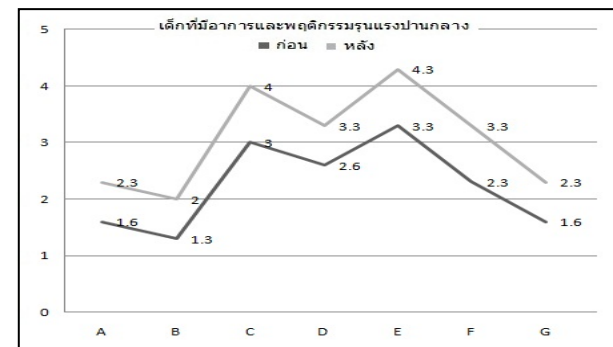


ภาพที่ 1: ค่าเฉลี่ยก่อนและหลัง เรื่องการบอกชื่อสี โทนสีแตกต่างกัน และโทนสีใกล้เคียงกันในกลุ่มอาการและพฤติกรรมรุนแรงปานกลาง

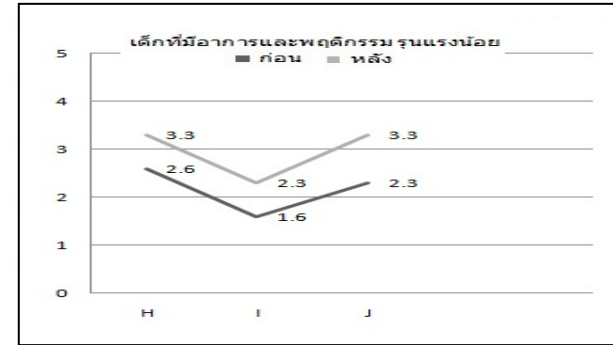


ภาพที่ 2: ค่าเฉลี่ยก่อนและหลัง เรื่องการบอกชื่อสี โทนสีแตกต่างกัน และโทนสีใกล้เคียงกันในกลุ่มอาการและพฤติกรรมรุนแรงน้อย

ภาพที่ 3 และภาพที่ 4 แสดงเรื่องการระบายสีโทนสีแตกต่างกัน และโทนสีใกล้เคียงกัน แสดงค่าเฉลี่ยก่อนและหลัง จากมุมมองผู้วิจัย ครูฝึก และผู้ปกครองเด็ก

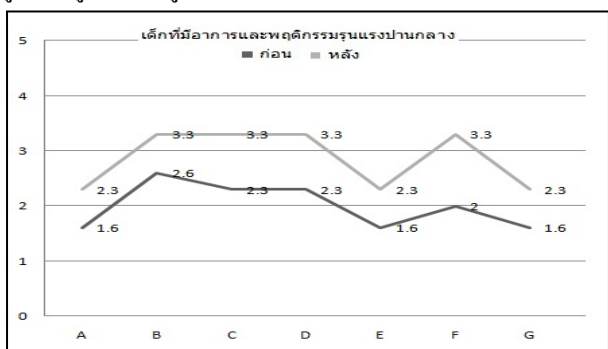


ภาพที่ 3: ค่าเฉลี่ยก่อนและหลัง เรื่องการระบายสี โทนสีแตกต่างกัน และโทนสีใกล้เคียงกันในกลุ่มอาการและพฤติกรรมรุนแรงปานกลาง

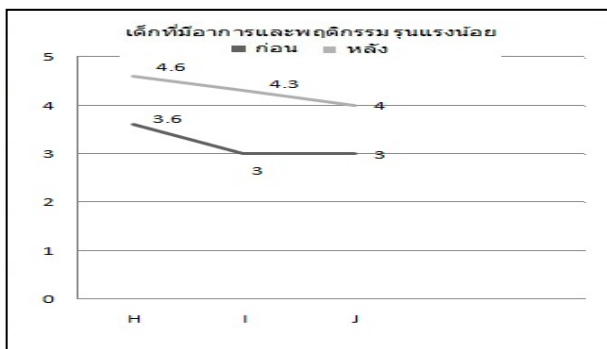


ภาพที่ 4: ค่าเฉลี่ยก่อนและหลัง เรื่องการระบายสี โทนสีแตกต่างกัน และโทนสีใกล้เคียงกันในกลุ่มอาการและพฤติกรรมรุนแรงน้อย

ภาพที่ 5 และภาพที่ 6 แสดงเรื่องการแยกสี โทนสีแตกต่างกัน และ โทนสีใกล้เคียงกัน แสดงค่าเฉลี่ยก่อนและหลัง จากมุมมอง ผู้วิจัย ครูฝึก และผู้ปกครองเด็ก



ภาพที่ 5: ค่าเฉลี่ยก่อนและหลัง เรื่องการแยกสี โทนสีแตกต่างกัน และ โทนสีใกล้เคียงกันในอากลุ่ม



ภาพที่ 6: ค่าเฉลี่ยก่อนและหลัง เรื่องการแยกสี โทนสีแตกต่างกัน และ โทนสีใกล้เคียงกันในกลุ่มอาการและพฤติกรรมรุนแรงน้อย

6. ข้อจำกัดในงานวิจัย

ในรายงานศึกษาวิจัยฉบับนี้ ไม่มีผลรายงานพฤติกรรมการเรียนรู้ของเด็กกลุ่มที่ 1 อาการและพฤติกรรมรุนแรงมาก เพราะเด็กในกลุ่มนี้ไม่สามารถใช้ชุดฝึกมัลติมีเดียเกม ศิลปะ ในการสร้างกระบวนการเรียนรู้ได้ และไม่มีผลการติดตามหลังจากเด็กผ่านการทดสอบไปแล้ว 7 สัปดาห์ของเด็กทั้งหมด ผู้วิจัยไม่สามารถทราบถึงผลพัฒนาการในการเรียนรู้เรื่องของเด็ก เพราะหากเด็กในกลุ่ม ข้ายศูนย์ฝึก และไม่ได้รับการพัฒนาอย่างต่อเนื่องจากผู้ปกครอง อาจส่งผลให้พัฒนาการในเรื่องสี ของเด็กไม่เกิดการพัฒนาอย่างต่อเนื่อง

7. อธิปราคาผล

ผลการศึกษาวิจัยที่มีผลต่อการสร้างพัฒนาการเด็กออทิสติกในศูนย์การศึกษาพิเศษจังหวัดหนองคาย พบว่า ความรู้ และเข้าใจในเรื่องของโรค จะทำให้สามารถดูแลเด็กออทิสติกได้อย่างมีประสิทธิภาพ[10] ซึ่งทำให้เกิดการเตรียม

ความพร้อมที่ดี ทำให้สามารถเผชิญ และแก้ไขปัญหาที่เกิดขึ้นได้อย่างเหมาะสม[11] และยังพบว่ามารดามีความคาดหวังและเชื่อว่าลูกของตนเองจะมีอาการที่ดีขึ้น[12]

หลังจากผ่านไป 7 สัปดาห์ พบว่า เด็กทั้ง 2 กลุ่มอาการมีพัฒนาการดีขึ้น โดยเด็กมีความสามารถในการบอกรหัสสีได้ดีขึ้น ความสามารถในการบังคับมือ(กล้ามเนื้อมัดเล็ก) ระบายสีโดยให้อยู่ในกรอบที่กำหนด เด็กทำได้ดีขึ้นเป็นลำดับ และยัง สามารถแยกความแตกต่างระหว่างสีเข้มและสีอ่อนได้ดีขึ้น เมื่อพิจารณาแยกกลุ่มอาการพบว่า เด็กในกลุ่มอาการและพฤติกรรมรุนแรงน้อย 1 คน พัฒนาการไม่ก้าวหน้า เมื่อตรวจสอบย้อนหลังจึงพบว่า ครอบครัวประกอบอาชีพ ทำนา ทำให้เด็กขาดการฝึกที่ต่อเนื่อง สอดคล้องกับงานวิจัยของ เพ็ญแข[13] กล่าวว่า การดูแลเด็กออทิสติกเป็นสิ่งที่ต้องกระทำด้วยความอดทน จริงจังอย่างต่อเนื่อง และต้องได้รับความช่วยเหลือจากบุคคลที่มีความรู้ และประสบการณ์จากหลายฝ่าย การนำเทคโนโลยีทางคอมพิวเตอร์ มาประยุกต์ใช้เป็นตัวกลางในการสื่อสารกับเด็กออทิสติกแทนการสอนนั้น สามารถสร้างแรงจูงใจ ดึงดูดความสนใจ และสอดแทรกความรู้ หรือเพิ่มทักษะ[14] โดยเทคโนโลยีสามารถช่วยอำนวยความสะดวกให้กับเด็กกลุ่มนี้ การนำเกมศิลปะมาช่วยเสริม ยังช่วยส่งเสริมการประสานกันระหว่างสายตาและมือของเด็ก ขณะจดจ่ออยู่กับการทำกิจกรรมที่เกิดขึ้นในเกม[15] ทำให้เด็กเกิดพัฒนาการ เพราะต้องคิดและตัดสินใจในการเลือกกิจกรรมที่เกิดขึ้นในเกม และกิจกรรมศิลปะช่วยทำให้เด็กมีสมาธิ ช่วยกล่อมเกลาคิดใจ ให้เด็กมีความละเอียดอ่อน ใจเย็น[8] และปัจจัยที่สำคัญอีกประการหนึ่ง ที่ค้นพบในงานวิจัยนี้ ผู้ปกครองต้องแบ่งเวลาในการเรียนรู้วิธีการฝึกเพื่อสามารถนำกลับไปฝึกต่อที่บ้านได้อย่างต่อเนื่อง เป็นการช่วยให้เด็กทำซ้ำๆ[16] เพื่อทำให้เด็กเกิดการเรียนรู้ จดจำ เพื่อทำให้เกิดความรู้ที่คงทน และเป็นสิ่งที่จำเป็นต่ออนาคตของตัวเด็กในการช่วยในการสร้างพัฒนาการต่อไปในอนาคต[17]

8. กิตติกรรมประกาศ

ขอขอบคุณ ผู้ปกครอง คณะครูฝึก และเด็กออทิสติกประจำศูนย์การศึกษาพิเศษประจำจังหวัดหนองคายทุกท่าน ที่สละเวลาให้ความร่วมมือกับทีมผู้วิจัย และขอบคุณผู้อำนวยการ (นางสาวธนากร ฤทธิผลวง) ศูนย์การศึกษาพิเศษประจำจังหวัด

หนองคายเป็นอย่างสูงที่อำนวยความสะดวกในการทำงานให้กับทีมผู้วิจัย

เอกสารอ้างอิง

- [1] รจนา ทรพรานนท์, เด็กออทิสติก...คำแนะนำสำหรับบิดามารดาและนักวิชาการ, กรุงเทพมหานคร: เกียรติธุรกิจ, 2527, หน้า 15.
- [2] Wing, L., Ealy Children Autism Oxford: Pregamon Press, 1976.
- [3] อุมพร ตรีภคสมบัติ, ช่วยลูกออทิสติก, กรุงเทพมหานคร: ศูนย์วิจัยและพัฒนาครอบครัว, 2550.
- [4] อรุณ จันทรวณิช, การพัฒนาเด็กออทิสติก, กรุงเทพมหานคร: กรมวิชาการ, ศูนย์พัฒนาหลักสูตร, 2543, หน้า 1.
- [5] สถาบันพัฒนาการเด็กราชชนรินทร์, การดูแลเด็กแบบบูรณาการ, เชียงใหม่: สถาบันพัฒนาการสถาบันพัฒนาการเด็กราชชนรินทร์ กรมสุขภาพจิต กระทรวงสาธารณสุข, 2546.
- [6] Archbold & Stawart B. J, Greelinck, M., & Harvath, T., Gadov, S., Kirschling, J. M., Miler, L. L, Brody, K., Schook, J, "Establishing partnerships with family caregivers: Local And cosmopolitan knowledge", *Journal of Gerontological Nursing*, 22(1), 1994.
- [7] N. Oudin O. Grynszpan, J.-C. Martin, "Towards a methodology for the design of humancomputer interfaces for persons with autism. ACM Crossroads", *Special Issue on Human-Computer Interaction*, Winter 2005.
- [8] รุสรวพร เต็มทอง และคณะ, "ผลของโปรแกรมดนตรีและศิลปะบำบัดที่มีต่อพฤติกรรมก้าวร้าวของเด็กบกพร่องทางพัฒนาการและสติปัญญาระดับรุนแรง เพศหญิง สถาบันราชานุกูล", สถาบันราชานุกูล, งานวิจัยและพัฒนานวัตกรรมองค์ความรู้, 2553.
- [9] นิภาพรรณ บุญชู, การศึกษาผลการสอนพูดโดยใช้โปรแกรม TEACCH สำหรับนักเรียนออทิสติก ชั้นเตรียมความพร้อม, หนองคาย: ศูนย์การศึกษาพิเศษประจำจังหวัดหนองคาย, 2554.
- [10] จันธิรา เตชะจักร, การดำเนินงานการให้บริการช่วยเหลือระยะแรกเริ่มแก่เด็กออทิสติกของสถาบันพัฒนาการเด็กราชชนรินทร์, วิทยานิพนธ์ปริญญา ศึกษาศาสตรมหาบัณฑิต สาขาวิชาการศึกษาพิเศษ คณะศึกษาศาสตร์ มหาวิทยาลัยเชียงใหม่, 2552.
- [11] สุดศิริ หิรัญขุนนะ, การพัฒนารูปแบบการดูแลสุขภาพที่บ้านและผู้ดูแลผู้ป่วยทางสังคม, วิทยานิพนธ์พยาบาลศาสตรดุษฎีบัณฑิต, บัณฑิตวิทยาลัย มหาวิทยาลัยมหิดล, 2541.
- [12] Matthem, J. Altire., & Silvia von Kluge, "Family Functioning and Coping Behaviors in Parents of Children with Autism", *Journal Child Fam Stud*, 18, pp. 83-92, 2009.
- [13] เพ็ญแข ล้อมีลา, รวมเรื่องวินิจฉัยโรคออทิสซึม, สมุทรปราการ, ช. แสงงานการพิมพ์, 2540.
- [14] S. Rong But Si. Stimulate Development [Online], Available : <http://ednet.kku.ac.th> [2012, October 01].
- [15] S. Bögels and C. Lamers. "The causal role of self-awareness in blushing-anxious, socially-anxious and social phobic individuals." *Behaviour Research and Therapy*. 40. 1367–1384, 2001.
- [16] Savner JL, Myles BS, Making visual supports work in the home and community: strategies for individuals with autism and Asperger syndrome, Autism Asperger Publishing, Shawnee Mission, 2000.
- [17] จิระพร ชะโน, การพัฒนารูปแบบการสอนสื่อความหมายเพื่อลดพฤติกรรมที่เป็นปัญหาในเด็กออทิสติกก่อนวัยเรียน, วิทยานิพนธ์ปริญญาศึกษาศาสตรดุษฎีบัณฑิต สาขาวิชาหลักสูตรและการสอน, มหาวิทยาลัยขอนแก่น, 2551.

การสร้างโจทย์คณิตศาสตร์ระดับชั้นประถมศึกษาปีที่ 1 ด้วยไวยากรณ์ไม่พึ่งบริบท: โปรแกรมประยุกต์บนแท็บเล็ตพีซีพร้อมระบบวิเคราะห์ผลการเรียนรู้แบบอัตโนมัติ

Creating a Math Class Grade 1 with CFG: Application on Tablet PCs with Automatic Analysis of Learning Outcomes

วชิระ แจ้งงาน (Wachira Khaengngan)¹ และ สุนิสา รีมเจริญ (Sunisa Rimcharoen)²

คณะวิทยาการสารสนเทศ มหาวิทยาลัยบูรพา
¹wajira.mit8@gmail.com, ²rsunisa@buu.ac.th

บทคัดย่อ

งานวิจัยนี้นำเสนอการใช้ไวยากรณ์ไม่พึ่งบริบทสำหรับสร้างแบบทดสอบวิชาคณิตศาสตร์แบบอัตโนมัติ เพื่อส่งเสริมพัฒนาการของนักเรียนชั้นประถมศึกษาปีที่ 1 โดยอ้างอิงเนื้อหาตามสาระการเรียนรู้จากหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พ.ศ. 2551 ผู้วิจัยได้ใช้ไวยากรณ์ดังกล่าวในการสร้างโจทย์และให้นักเรียนฝึกทำแบบฝึกหัดผ่านโปรแกรมประยุกต์บนแท็บเล็ตพีซี โดยมีชื่อโปรแกรมคือ Math4Kids และพัฒนาระบบสำหรับวิเคราะห์ผลเพื่อให้ผู้สอนสามารถหาจุดบกพร่องของนักเรียนและเข้าไปเสริมความรู้ให้กับนักเรียนได้ตรงจุดจากผลการทดลองกับนักเรียน 10 คน เป็นเวลาต่อเนื่อง 5 วัน พบว่า นักเรียนสามารถทำแบบฝึกหัดที่ยากขึ้นและสามารถทำแบบทดสอบได้ถูกต้องมากขึ้น ซึ่งผลจากการทำแบบทดสอบหลังใช้โปรแกรมประยุกต์บนแท็บเล็ตพีซีพบว่านักเรียนส่วนใหญ่ได้คะแนนมากกว่าการทำแบบทดสอบก่อนใช้โปรแกรมประยุกต์

คำสำคัญ: การสร้างโจทย์แบบฝึกหัดวิชาคณิตศาสตร์ แท็บเล็ตพีซี ไวยากรณ์ไม่พึ่งบริบท

Abstract

This paper presents using Context-Free Grammar to generate math problems automatically in order to help 1st grade students learn mathematics skills. The scope of the generated problems is based on Basic Education Core Curriculum (B.E.2551). The application program on tablet PC is developed for students to practice math exercises. The program is named Math4Kids. We also provide an analysis

tool as a web application so that teachers can notice the weaknesses of the students and give advice especially for each student. The experiments with 10 students for 5 consecutive days show that students can do the more difficult exercises and get more accurate results. In addition, after using tablet PC-based practice, most of the students have higher posttest score than the pretest.

Keywords: Math Problem Generation, Tablet PC, Context Free Grammar

1. บทนำ

คณิตศาสตร์เป็นวิชาหนึ่งที่มีความสำคัญทั้งในด้านชีวิตประจำวัน และความก้าวหน้าของวิทยาการต่างๆ อีกทั้งยังเป็นศาสตร์พื้นฐานของวิชาอื่น เช่น วิทยาศาสตร์ เศรษฐศาสตร์ เป็นต้น จากความสำคัญของวิชาคณิตศาสตร์ดังกล่าว กระทรวงศึกษาธิการ จึงกำหนดให้วิชาคณิตศาสตร์บรรจุอยู่ในการเรียนการสอนที่นอกจากมุ่งเน้นให้ผู้เรียนได้รับความรู้ความเข้าใจแล้วยังต้องให้ผู้เรียนได้ฝึกทักษะการคิดคำนวณ ซึ่งเมื่อจบการศึกษาขั้นพื้นฐาน 12 ปีแล้ว ผู้เรียนต้องมีความรู้ความเข้าใจในเนื้อหาคณิตศาสตร์และสามารถนำความรู้ทางคณิตศาสตร์ไปพัฒนาคุณภาพชีวิตและเป็นพื้นฐานการศึกษาในระดับสูงขึ้น [1]

วิธีหนึ่งในการฝึกทักษะการคำนวณทางคณิตศาสตร์ก็คือการฝึกทำแบบฝึกหัด การออกโจทย์นี้จะต้องสอดคล้องกับหลักสูตรและตัวชี้วัดของกลุ่มสาระการเรียนรู้ชั้นๆ การสร้างแบบฝึกหัดแต่ละครั้งผู้สอนต้องเสียเวลาออกโจทย์ ในบางครั้งจึงมีการนำโจทย์เก่ามาให้ให้นักเรียนทำ ซึ่งอาจทำให้ไม่เกิดความ

หลากหลายของตัวอย่าง โจทย์ ดังนั้น คณะผู้วิจัยจึงมีแนวคิดที่จะนำไวยากรณ์ไม่พึ่งบริบทมาช่วยในการสร้างแบบฝึกหัดอัตโนมัติ เพื่อที่จะช่วยลดเวลาในการสร้างแบบฝึกหัดและยังเพิ่มความหลากหลายของโจทย์คณิตศาสตร์อีกด้วย

ปัจจุบันการทำแบบทดสอบไม่ได้จำกัดแค่ในหนังสือแต่ยังมีรูปแบบอื่นๆ เช่น โปรแกรมประยุกต์ผ่านเว็บแอปพลิเคชัน หรือ โปรแกรมประยุกต์ที่ใช้งานบนอุปกรณ์เคลื่อนที่ เช่น สมาร์ทโฟน และแท็บเล็ตพีซี [2] ซึ่งเทคโนโลยีแท็บเล็ตพีซีนับได้ว่าเป็นสื่อที่กำลังได้รับความนิยม แม้แต่ภาครัฐฯ ยังได้กำหนดนโยบายการใช้เทคโนโลยีแท็บเล็ตพีซีกับนักเรียนระดับชั้นประถมศึกษาปีที่ 1 เพื่อใช้เป็นเครื่องมือสำหรับการเรียนการสอนยุคใหม่

ทั้งนี้ คณะผู้วิจัยได้สังเกตเห็นว่า การส่งเสริมทักษะวิชาคณิตศาสตร์โดยใช้เทคโนโลยีผ่านแท็บเล็ตพีซี จะช่วยพัฒนาทักษะและกระบวนการทางคณิตศาสตร์สำหรับนักเรียนชั้นประถมศึกษาปีที่ 1 และยังสามารถพัฒนาเป็นต้นแบบในการต่อยอดในอนาคต บทความนี้จึงได้เสนอวิธีสร้างโจทย์คณิตศาสตร์แบบอัตโนมัติโดยใช้ไวยากรณ์ไม่พึ่งบริบทบนแท็บเล็ตพีซี เพื่อผู้สร้างโจทย์ที่มีความหลากหลายและแบ่งระดับความยากของโจทย์ตามพัฒนาการของผู้เรียน ส่วนทางด้านผู้สอนก็สามารถติดตามผลการเรียนรู้ของนักเรียนแต่ละคนได้ผ่านโปรแกรมเว็บแอปพลิเคชัน ซึ่งจะแสดงผลพัฒนาการในการทำโจทย์แบบฝึกหัดของนักเรียน รวมทั้งบอกจุดบกพร่องของนักเรียนแต่ละคนเพื่อเป็นข้อมูลให้ผู้สอนสามารถเข้าไปเสริมความรู้ให้กับนักเรียนได้ตรงความต้องการ

2. งานวิจัยและทฤษฎีที่เกี่ยวข้อง

2.1 งานวิจัยที่เกี่ยวข้อง

บัวบาง บุญอยู่ [3] ได้กล่าวถึงการเรียนการสอนที่มีผู้เรียนเป็นศูนย์กลางว่า การเรียนที่มีประสิทธิภาพ คือ การเรียนที่เน้นให้ผู้เรียนได้ทำกิจกรรมมาก ซึ่งทำให้เด็กมีประสบการณ์ตรง และสามารถแสวงหาวิธีเรียนที่เหมาะสมกับตนเองได้

การนำแท็บเล็ตพีซีมาใช้เพื่อเป็นเครื่องมือในการเรียนการสอนเป็นวิธีหนึ่งที่จะใช้เป็นเครื่องมือในการส่งเสริมการเรียนรู้ของนักเรียน ซึ่งโจทย์ปัญหาคณิตศาสตร์ผู้แก้ปัญหาจะทำได้ต้องมีกระบวนการที่เหมาะสมซึ่งใช้ความรู้ ประสบการณ์ การวางแผน และการตัดสินใจประกอบกัน [4] ผลการวิจัยการใช้แท็บเล็ตพีซี

เป็นเครื่องมือประกอบการเรียนล้วนสรุปว่า การใช้เทคโนโลยีแท็บเล็ตพีซีช่วยส่งเสริม การเรียนรู้ของผู้เรียน ดังตัวอย่างต่อไปนี้

Becta ICT Research [5] ได้ใช้แท็บเล็ตพีซีประกอบการเรียนการสอนในโรงเรียนระดับประถมศึกษา พบว่า แท็บเล็ตพีซีช่วยเพิ่มแรงจูงใจของผู้เรียนและมีผลกระทบในทางบวกต่อผลสัมฤทธิ์ทางการเรียน รวมทั้งสนับสนุนให้เกิดการเรียนรู้ด้วยตนเองและส่งเสริมการเรียนรู้แบบมีส่วนร่วมของผู้เรียน

สุรศักดิ์ ปะเฮ [6] วิจัยการจัดการเรียนการสอนพบว่า การใช้แท็บเล็ตพีซีนั้นช่วยส่งเสริมให้มีการใช้เทคโนโลยีการเรียนการสอน และส่งเสริมให้มีการพัฒนาหลักสูตรหรือการจัดการเรียนการสอนที่มีเทคโนโลยีสารสนเทศเป็นส่วนประกอบมากขึ้น

2.2 กลุ่มสาระการเรียนรู้วิชาคณิตศาสตร์ระดับชั้นประถมศึกษาปีที่ 1

ตามหลักสูตรแกนกลางการศึกษา ขั้นพื้นฐาน พ.ศ.2551 ได้กำหนดสาระการเรียนรู้วิชาคณิตศาสตร์สำหรับนักเรียนชั้นประถมศึกษาปีที่ 1 โดยมีตัวชี้วัดและสาระการเรียนรู้แกนกลางไว้เป็นมาตรฐานสำหรับผู้สอน ทั้งนี้เพื่อเป็นการบริหารการจัดการศึกษาและพัฒนาผู้เรียนให้เกิดการเรียนรู้ตามธรรมชาติ และเต็มศักยภาพตามเจตนารมณ์ของพระราชบัญญัติการศึกษาแห่งชาติ พ.ศ. 2542 และฉบับปรับปรุง พ.ศ. 2545

สาระการเรียนรู้ของวิชาคณิตศาสตร์ที่เป็นมาตรฐานสำหรับนักเรียนชั้นประถมศึกษาปีที่ 1 มีวัตถุประสงค์เพื่อให้นักเรียนฝึกทักษะการคิดคำนวณ ดังหัวข้อต่อไปนี้

- 1) จำนวนนับ: จำนวนนับ 1 ถึง 100, จำนวนศูนย์, การอ่าน และการเขียนตัวเลขฮินดูอารบิกและตัวเลขไทย
- 2) การบวกและการลบ: การบวก การลบ โดยใช้เครื่องหมาย + และ -, การบวกที่ไม่มีตัวทด, การลบที่ไม่มีการกระจาย, การบวกลบระคน
- 3) โจทย์ปัญหา: โจทย์ปัญหาการบวก การลบ, การวัดความยาว ความสูง และระยะทางโดยใช้หน่วยที่ไม่ใช่หน่วยมาตรฐาน
- 4) การเปรียบเทียบความยาว, น้ำหนัก, การตวง, การชั่ง, ปริมาตร, ความจุ, เวลา

จากสาระการเรียนรู้ดังกล่าว ผู้สอนได้ให้สัมภาษณ์ว่า เนื้อหาที่นักเรียนส่วนใหญ่ทำไม่ได้คือการทำโจทย์ปัญหา ดังนั้นในงานวิจัยนี้ จึงมุ่งเน้นไปที่การสร้างโจทย์ ซึ่งมีทั้งโจทย์แบบง่ายที่

เป็นการบวกลบตัวเลขและโจทย์ปัญหาที่นักเรียนต้องอ่าน โจทย์ และตีความหมาย เช่น “มานะมีเงิน 10 บาท แม่ให้มาอีก 10 บาท รวมแล้วมานะมีเงินเท่าไร” ซึ่งวิธีการที่คณะผู้วิจัย จะใช้ในการสร้าง โจทย์ในลักษณะนี้คือ การสุ่มสร้าง โจทย์จากไวยากรณ์ไม่ พึ่งบริบท ซึ่งทฤษฎีของไวยากรณ์ดังกล่าวจะกล่าวถึงในหัวข้อ 2.3 และตัวอย่างของไวยากรณ์ที่จะใช้สร้าง โจทย์คณิตศาสตร์จะ แสดงในวิธีดำเนินการวิจัย

2.3 ไวยากรณ์ไม่พึ่งบริบท (Context-Free Grammar)

ไวยากรณ์ไม่พึ่งบริบท หมายถึง การบรรยายภาษาคำว ยกฎไวยากรณ์ ซึ่งประกอบด้วยส่วนสำคัญสี่ส่วน ดังนี้

$$G = (V, T, S, P)$$

โดยที่ V คือ เซตของตัวแปร (Non-Terminal)

T คือ เซตของตัวอักษร (Terminal) ซึ่งเป็นเซตที่มี สมาชิกไม่ซ้ำกับเซตของตัวแปร

S คือ ตัวแปรเริ่มต้น (Start Variable) โดยที่ $\epsilon \in V$

P คือ เซตจำของกฎไวยากรณ์ (Production Rules)

เป็นเซตของกฎซึ่งอยู่ในรูปของ

$$A \rightarrow \alpha \text{ เมื่อ } A \in V \text{ และ } \alpha \in (V \cup T)^*$$

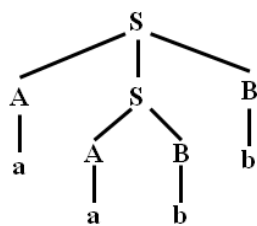
ตัวอย่างเช่น $\langle S \rangle \rightarrow \langle A \rangle \langle B \rangle$

$$\langle S \rangle \rightarrow \langle A \rangle \langle S \rangle \langle B \rangle$$

$$\langle A \rangle \rightarrow a$$

$$\langle B \rangle \rightarrow b$$

จากกฎข้างต้น นำมาเขียนเป็นต้นไม้แจงส่วน (Parse Tree) ของข้อความตัวอย่าง “aabb” ได้ดังรูปที่ 1



ภาพที่ 1 การต้นไม้แจงส่วน (Parse Tree)

3. วิธีดำเนินการวิจัย

3.1 ไวยากรณ์ไม่พึ่งบริบทสำหรับการสร้างแบบฝึกหัดวิชา คณิตศาสตร์

งานวิจัยนี้ได้ออกแบบไวยากรณ์ไม่พึ่งบริบทสำหรับสร้าง แบบทดสอบอัตโนมัติ ซึ่งคณะผู้วิจัยได้แบ่งลักษณะของ แบบทดสอบออกเป็นหมวด ดังนี้

1. โจทย์บวกแบบตัวเลข
2. โจทย์ลบแบบตัวเลข
3. โจทย์บวกแบบปัญหา
4. โจทย์ลบแบบปัญหา
5. โจทย์ระคนแบบบวก บวก
6. โจทย์ระคนแบบบวก ลบ
7. โจทย์ระคนแบบลบ บวก
8. โจทย์ระคนแบบลบ ลบ

โดยไวยากรณ์ที่ใช้ในการสร้าง โจทย์ปัญหาแสดงในรูปที่ 2

ตัวอย่าง โจทย์ที่สร้างได้จากไวยากรณ์ในรูปที่ 2 เช่น “พ่อมี มะม่วง 5 ผล กิน ไป 2 ผล เหลือกี่ผล” จะสามารถเขียนออกมา เป็นกฎไวยากรณ์ได้ดังนี้

$$\langle \text{Question} \rangle \rightarrow \langle \text{Subject} \rangle \langle \text{VP} \rangle$$

$$\langle \text{Subject} \rangle \rightarrow \langle \text{ทั่วไป} \rangle \langle \text{V} \rangle$$

$$\langle \text{ทั่วไป} \rangle \rightarrow \text{พ่อ}$$

$$\langle \text{V} \rangle \rightarrow \text{มี}$$

$$\langle \text{VP} \rangle \rightarrow \langle \text{SUB_จำนวน} \rangle$$

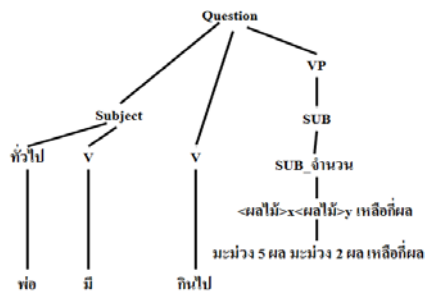
$$\langle \text{SUB_จำนวน} \rangle \rightarrow \langle \text{ผลไม้} \rangle X \langle \text{ผลไม้} \rangle Y \text{ เหลือกี่ผล}$$

$$\langle \text{ผลไม้} \rangle \rightarrow \text{มะม่วง}$$

ซึ่งมีวิธีการสร้างต้นไม้แจงส่วน (Parse Tree) ดังรูปที่ 3

3.2 การออกแบบโปรแกรมประยุกต์บนแท็บเล็ตพีซี

หลังจากที่คณะผู้วิจัยได้ออกแบบกฎการสร้างแบบฝึกหัด โดยใช้ไวยากรณ์ไม่พึ่งบริบทแล้ว ขั้นตอนต่อไปจึงทำการ ออกแบบ โปรแกรมประยุกต์บนแท็บเล็ตพีซีที่ใช้ ระบบปฏิบัติการแอนดรอยด์ ซึ่งคณะผู้วิจัยได้ศึกษาและพัฒนา แอปพลิเคชันด้วยภาษาจาวา (Java) และใช้โปรแกรม Eclipse เป็นเครื่องมือช่วยพัฒนาแอปพลิเคชัน และได้เลือกพัฒนาบน อุปกรณ์แท็บเล็ตพีซีสำหรับนักเรียนชั้นประถมศึกษาปีที่ 1



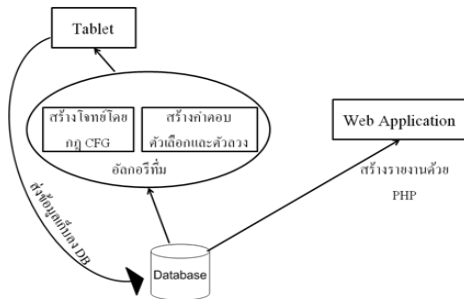
ภาพที่ 3 แสดงต้นไม้แจงส่วน

โปรแกรมพัฒนาขึ้นโดยใช้หลักการออกแบบคอมพิวเตอร์ มีลัทธิมีเดียผสมผสานกับการสร้างแบบทดสอบทางคณิตศาสตร์ มาประยุกต์ใช้ในการออกแบบโปรแกรมประยุกต์สำหรับสร้างแบบทดสอบ นั่นคือ ทำให้ง่ายต่อการใช้งาน คณะผู้วิจัยได้ออกแบบฟังก์ชันการทำงานของโปรแกรมประยุกต์สำหรับสร้างแบบทดสอบอัตโนมัติ โดยแบ่งหมวดหมู่ประเภทของแบบฝึกหัดตามเนื้อหา และคำนึงถึงความสามารถของนักเรียน โดยแบบฝึกหัดที่สร้างขึ้นมาจะเพิ่มความยากขึ้นตามทักษะของ

- <Question> → <Subject><VP>
- <Subject> → <เฉพาะ>|<ทั่วไป>
- <เฉพาะ> → “มานี่” | “มานะ” | “ปิติ” | “ซุใจ”
- <ทั่วไป> → “พ่อ” | “แม่” | “พี่” | “น้องชาย” | “น้องสาว” | “น้ำ”
- <VP> → <ADD>|<SUB>|<ADD_ADD>|<ADD_SUB>|<SUB_ADD>|<SUB_SUB>
- <ADD> → <ADD_น้ำหนัก>|<ADD_จำนวน>
- <ADD_น้ำหนัก> → <ผลไม้> x <ผลไม้> y รวมกันเท่าไร | <S_เฉพาะ> x <S_เฉพาะ> y รวมกันหนักเท่าไร | <เครื่องเขียน> x <เครื่องเขียน> y รวมกันเป็นเท่าไร | <สิ่งของ> x <สิ่งของ> y รวมกันเป็นเท่าไร
- <ADD_จำนวน> → <ผลไม้> x <ผลไม้> y รวมกันมีกี่ผล / ลูก <เครื่องเขียน> x <เครื่องเขียน> y รวมกันเป็นเท่าไร | <สิ่งของ> x <สิ่งของ> y รวมกันมีเท่าไร
- <SUB> → <SUB_น้ำหนัก> | <SUB_จำนวน>
- <SUB_น้ำหนัก> → <ผลไม้> x <ผลไม้> y เหลือเท่าไร | <เครื่องเขียน> x <เครื่องเขียน> y เหลือเท่าไร
- <SUB_จำนวน> → <ผลไม้> x <ผลไม้> y เหลือเท่าไร | <เครื่องเขียน> x <เครื่องเขียน> y เหลือกี่อัน
- <ADD_ADD> → <ADD_ADD_น้ำหนัก> | <ADD_ADD_จำนวน>
- <ADD_ADD_น้ำหนัก> → <ผลไม้> x <ผลไม้> y <ผลไม้> z รวมกันเป็นเท่าไร | <ผลไม้> x <ผลไม้> y <ผลไม้> z รวมเป็นกี่ผล
- <ADD_ADD_จำนวน> → <ผลไม้> x <ผลไม้> y <ผลไม้> z รวมทั้งหมดกี่ลูก | <ผลไม้> x <ผลไม้> y <ผลไม้> z รวมแล้วมีเท่าไร
- <ADD_SUB> → <ADD_SUB_จำนวน> | <ADD_SUB_น้ำหนัก>
- <ADD_SUB_จำนวน> → <ผลไม้> x <ผลไม้> y <ผลไม้> z เหลือเท่าไร | <ผลไม้> x <ผลไม้> y <ผลไม้> z เหลือกี่ผล
- <ADD_SUB_น้ำหนัก> → <ผลไม้> x <ผลไม้> y <ผลไม้> z เหลือเท่าไร | <ผลไม้> x <ผลไม้> y <ผลไม้> z มีเท่าไร
- <SUB_ADD> → <SUB_ADD_น้ำหนัก> | <SUB_ADD_จำนวน>
- <SUB_ADD_น้ำหนัก> → <ผลไม้> x <ผลไม้> y <ผลไม้> z เป็นเท่าไร
- <SUB_ADD_จำนวน> → <ผลไม้> x <ผลไม้> y <ผลไม้> z เป็นเท่าไร
- <SUB_SUB> → <SUB_SUB_จำนวน> | <SUB_SUB_น้ำหนัก>
- <SUB_SUB_จำนวน> → <ผลไม้> x <ผลไม้> y <ผลไม้> z เหลือเท่าไร | <เครื่องเขียน> x <เครื่องเขียน> y <เครื่องเขียน> z เหลือเท่าไร
- <SUB_น้ำหนัก> → <ผลไม้> x <ผลไม้> y <ผลไม้> z เหลือกี่กิโลกรัม | <เครื่องเขียน> x <เครื่องเขียน> y <เครื่องเขียน> z เหลือเท่าไร
- <ผลไม้> → “มะม่วง” | “มะละกอ” | “มังคุด” | “กล้วย” | “มะพร้าว” | “มะนาว” | “ส้ม” | “แตงโม” | “ส้มโอ” | “องุ่น”
- <เครื่องเขียน> → “ปากกา” | “ดินสอ” | “ยางลบ” | “สมุด” | “ไม้บรรทัด” | “กระดาษ”

ภาพที่ 2 ตัวอย่างไวยากรณ์ที่ใช้ในการสร้าง

ผู้เรียนแต่ละคน และผู้สอนสามารถตรวจสอบแบบทดสอบที่นักเรียนทำ และสามารถวิเคราะห์จุดอ่อนของนักเรียนได้อีกด้วย ภาพรวมของการทำงานของระบบแสดง ดังภาพที่ 4



ภาพที่ 4 ภาพรวมการทำงานของระบบ

จากรูปที่ 4 แสดงให้เห็นถึงภาพรวมการทำงานของระบบทั้งหมด โดยจะเริ่มจากนักเรียนใช้งานโปรแกรมประยุกต์ผ่านทางแท็บเล็ตพีซี โปรแกรมประยุกต์ก็จะส่งข้อมูลของแบบทดสอบของนักเรียนมาเก็บไว้ในฐานข้อมูล โดยที่ผู้สอนสามารถเรียกดูข้อมูลได้ผ่านทางเว็บแอปพลิเคชัน ซึ่งรายละเอียดของการวิเคราะห์ข้อมูลจะกล่าวต่อไปในหัวข้อที่ 3.3

สำหรับขั้นตอนการทำงานของระบบ เมื่อเริ่มทำการเข้าสู่ระบบ (Login) ข้อมูลการเข้าสู่ระบบก็ถูกส่งเข้าเก็บไว้ในฐานข้อมูล ในการสร้างโจทย์ เมื่อนักเรียนเข้าสู่การทำแบบฝึกหัด โปรแกรมจะทำการสร้างแบบฝึกหัดจากกฎของไวยากรณ์ที่คณะผู้วิจัยสร้างไว้ดังที่กล่าวไว้ในหัวข้อที่ 3.1 และ โปรแกรมจะไปสุ่มตัวเลือกจากฐานข้อมูล ซึ่งตัวเลือกที่เป็นตัวจริงจะมีจำนวนที่บวกเกินและลบออกจากคำตอบ หลังจากที่แบบฝึกหัดและตัวเลือกถูกสร้างขึ้นมาแล้ว นักเรียนจะทำการเลือกคำตอบ โปรแกรมจะส่งข้อมูลแบบฝึกหัด รูปแบบของตัวเลือกที่ทำการสุ่มออกมา รวมไปถึงคำตอบที่เลือก เข้าสู่ฐานข้อมูลผ่านเทคโนโลยี JSON โดยจะอ้างอิงจากรหัสสมาชิกและรหัสการเข้าสู่ระบบ ก่อนที่จะออกแบบฝึกหัดข้อใหม่ และเมื่อนักเรียนต้องการออกหรือเลิกใช้งาน โปรแกรม จะทำการออกจากระบบ ข้อมูลการออกจากระบบก็จะถูกส่งไปที่ฐานข้อมูล เพื่อใช้ในการหาเวลาที่อยู่ในระบบและการวิเคราะห์ข้อมูลในขั้นตอนต่อไป

3.3 ระบบวิเคราะห์ข้อมูล

หลังจากที่ได้ข้อมูลจากการทำแบบฝึกหัดของนักเรียนแล้ว คณะผู้วิจัยได้เพิ่มการวิเคราะห์ข้อมูลเพื่อให้ผู้สอนได้ติดตาม

พัฒนาการในการทำแบบทดสอบของนักเรียนผ่านทางเว็บแอปพลิเคชัน ซึ่งคุณสมบัติในส่วนนี้แตกต่างจากโปรแกรมฝึกทำ

โจทย์คณิตศาสตร์บนแท็บเล็ตอื่น โดยระบบที่ผู้วิจัยได้พัฒนาขึ้นจะนำข้อมูลที่ได้ออกมาวิเคราะห์ ดังนี้

- 1) เวลาที่นักเรียนใช้ทำแบบทดสอบในแต่ละครั้ง (การเข้าสู่ระบบ)
- 2) เวลาที่ใช้ในการวิเคราะห์แบบทดสอบ ระยะเวลาในการทำแบบทดสอบต่อข้อ (ข้อต่อวินาที)
- 3) รูปแบบของคำถามและคำตอบที่นักเรียนตอบผิด
- 4) ร้อยละความถูกต้องในการทำข้อสอบ

โดยระบบวิเคราะห์สามารถดูข้อมูลของทั้งกลุ่มหรือเลือกดูเฉพาะรายบุคคลได้

4. ผลการทดลอง

4.1 การทดสอบการสร้างโจทย์ด้วย CFG

หลังจากทดสอบการสร้างโจทย์โดยใช้ CFG ปรากฏว่าโจทย์ที่ถูกสร้างออกมาอยู่ในเกณฑ์ดี นักเรียนสามารถเข้าใจโจทย์และสามารถทำโจทย์ได้ถูกต้อง

4.2 การใช้งานแอปพลิเคชัน

หลังจากที่คณะผู้วิจัยพัฒนาโปรแกรมประยุกต์บนแท็บเล็ตพีซีและระบบวิเคราะห์ข้อมูลบนเว็บแอปพลิเคชันแล้ว จึงได้นำโปรแกรมประยุกต์ไปติดตั้งบนแท็บเล็ตพีซีและให้นักเรียนกลุ่มตัวอย่างจำนวน 10 คน ซึ่งนักเรียนกลุ่มนี้เก่ง-อ่อนคณิตศาสตร์ ละครัน โดยครูประจำชั้นเป็นคนเลือกมี เด็กเก่ง 5 คน เด็กอ่อน 5 คน โดยใช้เวลาดังสิ้น 5 วันต่อเนื่อง โดยก่อนการทดลองคณะผู้วิจัยได้ให้นักเรียนกลุ่มตัวอย่างทำแบบทดสอบก่อนใช้โปรแกรมประยุกต์และเมื่อใช้งานโปรแกรมบนแท็บเล็ตพีซีไป 5 วันก็ให้นักเรียนคนเดิมทำแบบฝึกหัดชุดเดิมหลังการใช้งานอีกครั้ง ซึ่งคำถามก่อนและหลังการทดลองจะเป็นโจทย์ที่นักเรียนได้เรียนในวิชาคณิตศาสตร์ ผลการทดลองพบว่านักเรียน 7 คนได้คะแนนมากขึ้น และอีก 3 คนได้คะแนนเท่าเดิม โดยนักเรียนทำคะแนนได้มากขึ้นเฉลี่ย 11 เปอร์เซ็นต์

เมื่อวิเคราะห์ข้อมูลโจทย์แต่ละประเภทที่นักเรียนกลุ่มทดลองได้ทำ พบว่านักเรียนกลุ่มนี้สามารถทำโจทย์ในประเภทของคำถามที่ 1-5 ได้ถูกต้อง แต่เมื่อเป็นโจทย์ที่ต้องมีการวิเคราะห์คำถาม 2 ครั้ง นักเรียนจะทำได้น้อย (คำถามที่ 6 และ 7) หรือไม่ยอมทำเลยเพราะคิดว่า ยากเกินไป (คำถามที่ 8) ซึ่งรายละเอียดแสดงดังตารางที่ 1

ตารางที่ 1 สรุปประเภทคำถามและความถูกต้อง

ประเภทของคำถาม	ข้อสอบที่ ทำทั้งหมด	ข้อสอบ ที่ทำถูก	ถูกต้อง (%)
1. โจทย์บวกแบบตัวเลข	571 ข้อ	530 ข้อ	92.82%
2. ลบแบบตัวเลข	294 ข้อ	243 ข้อ	82.65%
3. ลบแบบโจทย์	200 ข้อ	168 ข้อ	84.00%
4. บวกแบบโจทย์	349 ข้อ	314 ข้อ	89.97%
5. ระคนแบบบวก บวก	362 ข้อ	295 ข้อ	81.49%
6. ระคนแบบบวก ลบ	392 ข้อ	291 ข้อ	74.23%
7. ระคนแบบลบ บวก	107 ข้อ	61 ข้อ	57.01%
8. ระคนแบบลบ ลบ	47 ข้อ	22 ข้อ	46.81%

ส่วนผลการวิเคราะห์ข้อมูลรายบุคคล เพื่อพัฒนาการในการคิด วิเคราะห์แบบฝึกหัด แสดงผลข้อมูลในตารางที่ 2

ตารางที่ 2 พัฒนาการในการทำแบบทดสอบในแต่ละครั้ง

ครั้งที่	เวลาที่อยู่ในระบบ (นาทีก)	จำนวนข้อที่ทำ	จำนวนข้อที่ทำได้	ถูกต้อง (%)
1	14:23	37	36	97.30%
2	07:26	29	28	96.55%
3	07:48	19	19	100.00%
4	07:13	21	21	100.00%
5	06:33	17	17	100.00%

ข้อมูลที่เป็นประโยชน์ต่อผู้สอนอีกอย่างหนึ่งคือระบบแสดงข้อมูลเพื่อประกอบการตัดสินใจของผู้สอน โดยระบบจะแสดงข้อมูลของแบบฝึกหัดที่นักเรียนได้ทำไป ตัวลวงคำตอบทั้ง 4 ข้อ รวมถึงคำตอบที่นักเรียนเลือก ดังแสดงในตารางที่ 3

ตารางที่ 3 แบบฝึกหัดที่นักเรียนตอบผิด

คำถาม	1	2	3	4	คำตอบ ที่เลือก
$(3 - 3) + 2 = ?$	0	2	10	1	4
$(3 + 5) + 2 = ?$	0	-2	-10	-1	4
$(3 + 6) - 2 = ?$	1	2	-2	0	2
แม่มีมะนาว 7 ผล ซื้อมาเพิ่ม 3 ผลรวมแล้วมีเท่าไร?	3	0	-3	1	3
มานีส้มมะนาว 2 ผล เพื่อนให้ 6 ผล รวมกันเป็นเท่าไร?	1	1	-3	0	3

จากข้อมูลในตารางที่ 3 ขอยกตัวอย่างบางข้อ เช่น “แม่มีมะนาว 7 ผล ซื้อมาเพิ่ม 3 ผลรวมแล้วมีเท่าไร?” ในข้อนี้ นักเรียนเลือก

คำตอบข้อ 3 (-3) แสดงว่า นักเรียนยังวิเคราะห์โจทย์ปัญหาไม่ได้ โดยนักเรียนอาจคิดว่า โจทย์ข้อนี้เป็นโจทย์แบบลบ ตัวอย่างข้อ “ $(3 + 6) - 2 = ?$ ” สำหรับข้อนี้ นักเรียนเลือกคำตอบข้อ 2 (2) แสดงว่า นักเรียนยังทำโจทย์ระคนไม่ได้ ซึ่งวิเคราะห์ได้ว่า นักเรียนหาผลลัพธ์จากตัวเลขในวงเล็บ $(3+6)$ และเลือกคำตอบโดยไม่เอาผลลัพธ์ในวงเล็บมาลบกับตัวเลขข้างนอกวงเล็บ (2) นักเรียนจึงเลือกคำตอบที่เป็นตัวลวงซึ่งสามารถอธิบายรูปแบบของตัวลวงที่นักเรียนทำได้

5. สรุป

งานวิจัยนี้นำเสนอโปรแกรม Math4kids ซึ่งเป็นโปรแกรมประยุกต์บนแท็บเล็ตพีซีที่ใช้ไวยากรณ์ไม่พึ่งบริบทในการสร้างแบบฝึกหัดอัตโนมัติ เพื่อส่งเสริมพัฒนาการด้านทักษะการคิดวิเคราะห์ จำนวนตัวเลขวิชาคณิตศาสตร์สำหรับชั้นประถมศึกษาปีที่ 1 และยังมีระบบวิเคราะห์ข้อมูลผ่านทางเว็บแอปพลิเคชันที่จะช่วยให้ผู้สอนเสริมความรู้ให้กับเด็กได้ตรงจุด และยังสามารถพัฒนาต่อยอดในอนาคตได้อีกด้วย

เอกสารอ้างอิง

- [1] กระทรวงศึกษาธิการ. หลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551. กรุงเทพฯ: โรงพิมพ์ชุมนุมสหกรณ์การเกษตรแห่งประเทศไทย, 2551.
- [2] ไพฑูรย์ ศรีฟ้า. เปิดโลก Tablet สู่ทิศทางการวิจัยด้านเทคโนโลยีและสื่อสารการศึกษา: จากแนวคิดสู่กระบวนการปฏิบัติ. มหาวิทยาลัยเกษตรศาสตร์, 2554.
- [3] บัวบาง บุญอยู่. การเปรียบเทียบผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ เรื่อง โจทย์ปัญหาระคนของนักเรียนชั้นประถมศึกษาปีที่ 2 มีเรียนด้วยวิธีสอนโดยนักเรียนเป็นศูนย์กลางและวิธีสอนแบบปกติ. ชลบุรี: มหาวิทยาลัยบูรพา, 2543.
- [4] Ander, K.B., and R.E. Pingre. Problem Solving in Mathematics. The National Council of Teachers of Mathematics. New York, 1973.
- [5] Becta ICT Research (2005). Tablet PCs in School Study Report. Retrieved September 15, 2012 from <http://oro.open.ac.uk/6407/1/>.
- [6] สุรศักดิ์ ปะเฮ. เอกสารประกอบการบรรยาย “แท็บเล็ต (Tablet) กับการจัดการศึกษาสำหรับนักเรียนในยุคศตวรรษที่ 21”. นครปฐม: ภาควิชาเทคโนโลยีการศึกษา มหาวิทยาลัยศิลปากร, 2555.

การพัฒนาสื่อการเรียนรู้คำศัพท์ภาษาอังกฤษจากนิทานเรื่องหนูน้อยหมวกแดงบนแท็บเล็ต

The Development of Learning English Vocabulary with Little Red Riding Hood on Tablet

ณัฐวดี หงษ์บุญมี (Nattavadee Hongboonmee)¹ และพิชิตรา ศิริวัฒน์ (Phichitra Siriwat)²
ภาควิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยนครสวรรค์
nattavadeeho@nu.ac.th¹

บทคัดย่อ

ภาษาอังกฤษเป็นภาษาสากลที่ใช้กันทั่วโลก และมีความจำเป็นอย่างมากในอนาคต จึงควรมีการเพิ่มทักษะภาษาอังกฤษให้กับเด็กให้มากขึ้น เพื่อการใช้ภาษาอังกฤษอย่างมีประสิทธิภาพในอนาคต คณะผู้ศึกษาจึงพัฒนาสื่อการเรียนรู้คำศัพท์ภาษาอังกฤษจากนิทานเรื่องหนูน้อยหมวกแดง โดยมีจุดมุ่งหมายดังนี้ 1) เพื่อส่งเสริมทักษะคำศัพท์ภาษาอังกฤษให้กับเด็ก 2) เพื่อศึกษาและพัฒนาสื่อการเรียนรู้บนแท็บเล็ตระบบแอนดรอยด์ ซึ่งทั้งหมดนี้มีภาพประกอบสีสันสวยงามและมีเสียงประกอบสร้างความน่าสนใจ และดึงดูดใจให้เกิดความอยากเรียนมากขึ้น ทำให้เป็นประโยชน์ต่อการเสริมทักษะคำศัพท์ภาษาอังกฤษ ให้กับเด็กไปพร้อมกับความสนุกสนาน และผลจากการประเมินความพึงพอใจของผู้ใช้งาน ได้ค่า 4.08 จากผู้ใช้ 52 คน และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.04 ซึ่งถือว่าการเรียนรู้คำศัพท์ภาษาอังกฤษจากนิทาน เรื่องหนูน้อยหมวกแดงบนแท็บเล็ตนี้ มีความพึงพอใจอยู่ในระดับดี

คำสำคัญ: คำศัพท์ภาษาอังกฤษ แอนิเมชัน แท็บเล็ต

Abstract

English is an international language used in the world and the necessary in the future. Therefore, there should be more English language skills to children, to use English effectively in the future. The study of development of learning English vocabulary with Little Red Riding Hood on tablet, the aim is 1) to promote English vocabulary skills to children 2) to study and develop the learning materials on the tablet android

system. All of which are illustrated, and a colourful assembly. Create interesting and attractive to want to study more. Be beneficial to learning English vocabulary to children along with fun. The assessment of user satisfaction is 4.08 for 52 users and standard deviation is 0.04. Which is considered the learning English vocabulary with Little Red Riding Hood on tablet satisfaction at a good level.

Keywords: English Vocabulary, Animation, Tablet

1. บทนำ

ในปัจจุบันมีการก้าวหน้าของเทคโนโลยีเป็นอย่างมาก และได้เข้ามามีบทบาททางการศึกษาของเด็กไทยมากขึ้น โดยเฉพาะอย่างยิ่งแท็บเล็ต จากนโยบาย One Tablet PC Per Child ของรัฐบาล ซึ่งนอกจากจะใช้ในการศึกษาตามหลักสูตรแล้ว ยังสามารถนำไปใช้ในการศึกษาเพิ่มเติม จากสื่อประเภทต่างๆ ได้เป็นอย่างดี และภาษาอังกฤษเป็นภาษาสากลที่ใช้สื่อสารกันทั่วโลก จึงควรให้ความสำคัญในการศึกษาภาษาอังกฤษ เนื่องจากในอนาคตอันใกล้ ประเทศไทยกำลังจะเข้าสู่ประชาคมอาเซียน ทำให้ภาษาอังกฤษมีความสำคัญมากยิ่งขึ้น แต่เนื่องจากประเทศไทย ยังไม่มีความพร้อมทางด้านภาษาอังกฤษมากเพียงพอ [1] จึงควรศึกษาและปลูกฝังภาษาอังกฤษให้กับเด็กไทยให้มากยิ่งขึ้น แต่การศึกษาคำศัพท์ภาษาอังกฤษ เป็นเรื่องที่ไม่น่าสนใจและยากสำหรับเด็ก ทำให้เด็กไม่มีความสนใจที่จะศึกษา คณะผู้วิจัยจึงเกิดแนวคิดว่าจะหาวิธีการศึกษาที่ทำให้เด็กมีความสนใจ และอยากที่จะศึกษาคำศัพท์ภาษาอังกฤษมากขึ้นโดยการใช้สื่อต่างๆ

จากที่ได้กล่าวมาข้างต้น คณะผู้วิจัยจึงได้คิดพัฒนาแอปพลิเคชันสื่อการเรียนรู้คำศัพท์ภาษาอังกฤษจากนิทาน เรื่องหนูน้อยหมวกแดงบนแท็บเล็ต เพื่อเสริมทักษะคำศัพท์ภาษาอังกฤษให้กับเด็ก โดยสื่อการเรียนรู้คำศัพท์ภาษาอังกฤษนี้ มีการสอดแทรกนิทานเรื่องหนูน้อยหมวกแดง เป็นภาษาอังกฤษและนิทานภาษาไทย คำศัพท์ภาษาอังกฤษแบ่งเป็นหมวดต่างๆ 7 หมวด แบบฝึกหัดแบ่งออกเป็น 7 หมวด และเกมให้เล่นเพื่อความสนุกสนานในระหว่างการเรียนรู้ เพื่อให้มีความสนุกสนานไปพร้อมกับการศึกษาคำศัพท์ และดึงดูดใจให้เด็กมีความสนใจที่จะศึกษามากขึ้น นำไปสู่การใช้ภาษาอังกฤษอย่างมีประสิทธิภาพต่อไป

2. วรรณกรรมที่เกี่ยวข้อง

ในการศึกษาค้นคว้าพัฒนาสื่อการเรียนรู้คำศัพท์ภาษาอังกฤษบนแท็บเล็ต คณะผู้วิจัยได้ทำการค้นคว้าจากเอกสาร งานวิจัยที่เกี่ยวข้อง ดังต่อไปนี้

2.1 แอนดรอยด์ (Android)

Android [2] คือ ระบบปฏิบัติการสำหรับอุปกรณ์พกพาเคลื่อนที่เช่น โทรศัพท์มือถือ แท็บเล็ต มีพื้นฐานอยู่บนระบบลินุกซ์เคอร์เนลซึ่งใช้ Android SDK (Software Development Kit) เป็นเครื่องมือสำหรับการพัฒนาแอปพลิเคชันบนระบบปฏิบัติการแอนดรอยด์ โดยใช้ภาษาจาวาในการพัฒนามีลักษณะเป็นระบบปฏิบัติการแบบ open source นักพัฒนาสามารถศึกษาและใช้งาน source code ของระบบปฏิบัติการได้

แอนดรอยด์เริ่มพัฒนาโดยบริษัทแอนดรอยด์ [3] ต่อมาได้ผนวกเข้ากับบริษัท Google และได้มีการร่วมมือกันระหว่างบริษัทต่างๆหลายบริษัทในการพัฒนาระบบแอนดรอยด์

2.2 แท็บเล็ต (Tablet)

แท็บเล็ต [4] คือ เครื่องคอมพิวเตอร์ส่วนบุคคลที่สามารถพกพาได้ และใช้หน้าจอสัมผัสในการทำงานเป็นอันดับแรก ออกแบบให้สามารถทำงานได้ด้วยตัวเอง แท็บเล็ตเริ่มเป็นที่นิยมมากขึ้น เมื่อบริษัทแอปเปิลเปิดตัว iPad เมื่อปีคริสตศักราช 2010 แท็บเล็ตพีซีไม่เหมือนกับคอมพิวเตอร์ตั้งโต๊ะหรือ Laptops ตรงที่ไม่มีแป้นพิมพ์ในการทำงาน แต่ใช้แป้นพิมพ์

เสมือนจริงในการทำงานแทน (มีแป้นพิมพ์ปรากฏบนหน้าจอ ใช้การสัมผัสในการพิมพ์) แท็บเล็ตพีซีทุกเครื่องจะมีอุปกรณ์ไร้สายสำหรับการเชื่อมต่ออินเทอร์เน็ตและระบบเครือข่ายภายใน

2.3 ActionScript 3.0

ActionScript 3.0 [5] เป็นภาษาที่เพิ่มความสามารถเชิงโต้ตอบให้กับการแสดงผลเช่น การทำเกม การสร้างฟอร์มสำหรับผู้ใช้ การเพิ่มเทคนิคภาพเคลื่อนไหว เป็นภาษาแบบ Script Language [6] มีลักษณะในเชิงออบเจ็กต์ (Object-Oriented Programming) และถูกประมวลผลโดย virtual machine ที่มาพร้อมกับ Flash Player 9 ซึ่งใช้ประมวลผลเฉพาะ ActionScript 3.0 เท่านั้น

2.4 นิทานเรื่องหนูน้อยหมวกแดง

นิทานเรื่องหนูน้อยหมวกแดง เป็นนิทานพื้นบ้านที่มีการเล่ากันปากต่อปากอย่างแพร่หลายในยุโรป เนื้อเรื่องมีหลากหลายรูปแบบ และถูกเปลี่ยนแปลงเนื้อเรื่องหลายครั้ง จนกระทั่งมีเนื้อหาอย่างในปัจจุบัน มีการสอดแทรกข้อคิดต่างๆ ได้แก่ การไม่เชื่อฟังผู้ใหญ่ และการไว้ใจคนแปลกหน้า

เนื้อเรื่องนิทาน [7] จะกล่าวถึง เด็กผู้หญิงที่ใส่ชุดคลุมสีแดง ถือตะกร้าอาหาร ไปเยี่ยมคุณยายที่ป่วยอยู่ โดยที่คุณแม่ของเธอได้เตือนว่าอย่าเถลไถล และอย่าพูดคุยกับคนแปลกหน้า แต่เธอก็หนีออกนอกเหนือคำสั่งของคุณแม่ และถูกหมาป่าใช้อุบายหลอกลวงจนทำให้ตัวเธอเองและคุณยายต้องเดือดร้อน โชคยังดีที่มีนายพรายมาช่วยเอาไว้และไล่หมาป่าไป

2.5 นิยามของเด็ก

เด็ก (child) [8] ตามพจนานุกรมฉบับราชบัณฑิตยสถาน พ.ศ.2542 นิยามว่าเด็ก คือคนที่มีอายุน้อยหรือผู้เยาว์ตามประมวลกฎหมายแพ่งและพาณิชย์ คือผู้ซึ่งอายุไม่ครบ 18 ปี บริบูรณ์หรือบุคคลอายุเกิน 7 ปีแต่ยังไม่เกิน 14 ปีบริบูรณ์

ในทางการแพทย์ เด็กคือผู้ที่อายุไม่เกิน 14 ปีบริบูรณ์หรือเมื่อใช้คำนำหน้าชื่อว่า เด็กชาย เด็กหญิง

2.6 งานวิจัยเกี่ยวกับนิทานและการศึกษาคำศัพท์

ภาษาอังกฤษ

สุมนทาวงษ์สวัสดิ์ [9] ได้ทำการวิจัยเรื่อง การใช้นิทาน

เป็นสื่อในการเรียนรู้ เพื่อพัฒนาความสามารถด้านการฟังภาษาอังกฤษของนักเรียนชั้นประถมศึกษาปีที่ 5 โรงเรียนประถมนนทบุรี กรุงเทพมหานคร จากผลการวิจัยพบว่า การใช้นิทานเป็นสื่อในการเรียนรู้เพื่อพัฒนาความสามารถด้านการฟังภาษาอังกฤษของนักเรียน พบว่า นักเรียนกลุ่มตัวอย่างมีความสามารถด้านการฟังภาษาอังกฤษสูงขึ้น และนักเรียนมีการพัฒนาด้านการฟังภาษาอังกฤษเพิ่มขึ้นทุกระยะ จากระดับปานกลางจนถึงมาก ภาพประกอบช่วยให้เข้าใจนิทานขณะฟังได้ดีขึ้น ส่วนการใช้เสียงประกอบในนิทาน ซึ่งมีเสียงเพลง เสียงดนตรีและเสียงประกอบ (effect) ทำให้บทเรียนมีความสมจริงตามบริบทของเนื้อเรื่องในนิทาน ส่งผลให้นักเรียนมีอารมณ์คล้อยตามเหตุการณ์ของเรื่อง รู้สึกตื่นเต้นเร้าใจ และเข้าใจเนื้อเรื่องมากขึ้น

3. วิธีการดำเนินการวิจัย

ในการศึกษาและพัฒนา สื่อการเรียนรู้คำศัพท์ภาษาอังกฤษจากนิทานเรื่องหนูน้อยหมวกแดงบนแท็บเล็ต คณะผู้ศึกษาได้ดำเนินการศึกษา และพัฒนาแอปพลิเคชันบนระบบแอนดรอยด์ เพื่อเป็นสื่อการเรียนรู้โดยทำเป็นนิทานภาษาไทย นิทานภาษาอังกฤษ คำศัพท์ภาษาอังกฤษ บททดสอบและเกม เพื่อให้ได้รับความรู้พร้อมความสนุกสนาน และเครื่องมือที่ใช้ในการดำเนินงานมีดังต่อไปนี้

3.1 เครื่องมือที่ใช้ในการดำเนินงาน

Software

- Adobe Flash CS 6 สำหรับวาดภาพประกอบสื่อการเรียนรู้และพัฒนาสื่อการเรียนรู้

- Cool Edit Pro 2.1 สำหรับบันทึก และตัดต่อเสียงที่ใช้ในสื่อการเรียนรู้

- Adobe Photoshop CS5 สำหรับทำไอคอนของสื่อการเรียนรู้

Hardware

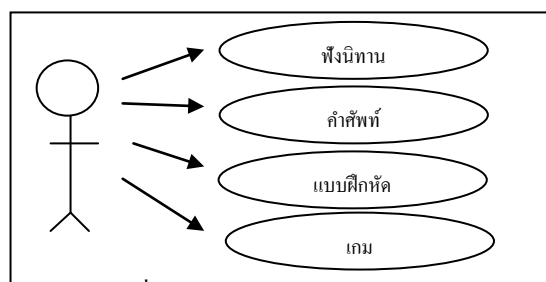
- เครื่องคอมพิวเตอร์ หน่วยประมวลผลกลาง CPU Intel Core i5 หน่วยความจำแรม 4 GB DDR RAM ฮาร์ดดิสก์ขนาด 500 GB

- Tablet รุ่น Lenovo หน่วยประมวลผลความเร็ว 1.2 GHz

ระบบปฏิบัติการ Android 4.1 หน้าจอแสดงผลขนาด 7 นิ้ว และความละเอียดหน้าจอ 600 x 1024 พิกเซล

3.2 การออกแบบและพัฒนาโปรแกรม

คณะผู้ศึกษาได้ทำการศึกษาข้อมูลที่ได้จากการค้นคว้า โดยกำหนดกลุ่มเป้าหมายหลัก ของการพัฒนาสื่อการเรียนรู้คือเด็กในช่วงอายุระหว่าง 10-14 ปี เนื่องจากช่วงอายุดังกล่าว มีความเหมาะสมสำหรับการเรียนรู้ภาษาอังกฤษ [10] และทำการวิเคราะห์ออกแบบโปรแกรม ได้ขอบเขตการทำงานของโปรแกรมดัง ภาพที่ 1



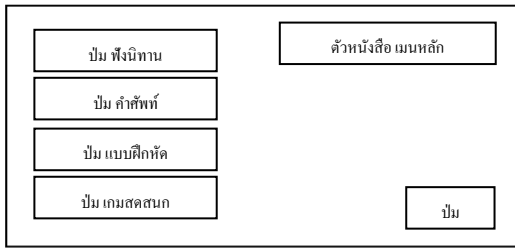
ภาพที่ 1: ขอบเขตการทำงานของโปรแกรม

โดยจากการวิเคราะห์และออกแบบโปรแกรม ได้แบ่งการทำงานของโปรแกรมออกเป็น 4 ส่วนหลัก คือ ส่วนของฟังนิทาน ซึ่งแบ่งเป็น การฟังนิทานภาษาอังกฤษและการฟังนิทานภาษาไทย ส่วนของคำศัพท์แบ่งออกเป็น 7 หมวด คือ คำศัพท์ในหมวดของ colors, food, animals, flowers, furniture, tools และ face ส่วนของแบบฝึกหัดแบ่งออกเป็น 7 หมวดเหมือนกับในหมวดคำศัพท์ ในแต่ละหมวดจะมีแบบฝึกหัดทั้งหมด 10 ข้อ และสุดท้ายเป็นเกมสุดสนุก ชื่อว่า เกมทายสีใครเอ่ย?

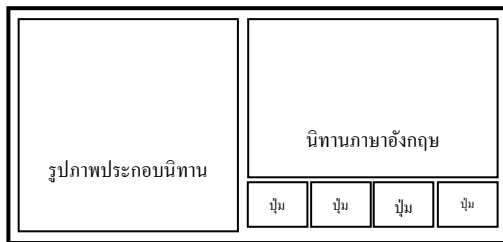
เมื่อวิเคราะห์และออกแบบโปรแกรมแล้ว ขั้นตอนต่อไปของการดำเนินงานมีดังต่อไปนี้

3.2.1 ออกแบบโครงสร้างหน้าจอโปรแกรม

เมื่อได้ขอบเขตการทำงานหลักของระบบแล้ว ขั้นตอนต่อไปเป็นการออกแบบโครงสร้างหน้าจอระบบทั้งหมด โดยตัวอย่างการออกแบบโครงสร้างหน้าจอโปรแกรม แสดงตัวอย่างภาพที่ 2 และภาพที่ 3



ภาพที่ 2: โครงร่างหน้าจอเมนูหลัก



ภาพที่ 3: โครงร่างหน้าจอนิทานภาษาอังกฤษ

3.2.2 ออกแบบตัวละครในนิทาน



ภาพที่ 4: หนูน้อยหมวกแดง

การออกแบบตัวละคร เช่น หนูน้อยหมวกแดง เด็กผู้หญิงใส่ชุดคลุมสีแดง ถือตะกร้าไปเยี่ยมคุณยายที่กำลังป่วยอยู่ และตัวละครอื่นๆ เช่น หม่าป้า คุณยาย คุณแม่ นายพราน

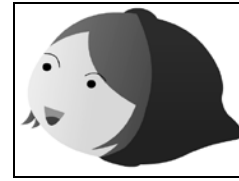
3.2.3 ขั้นตอนการวาดภาพประกอบ



ภาพที่ 5: ภาพลายเส้นหนูน้อยหมวกแดง

วิธีการวาดภาพประกอบมีขั้นตอนดังนี้ วาดรูปลายเส้นและสแกนภาพลงคอมพิวเตอร์ import รูปลงโปรแกรม Adobe Flash CS6 และสร้างรูปโดยใช้เครื่องมือ Line Tool ,Oval Tool และ Selection Tool ลงสีโดยใช้เครื่องมือ Paint Bucket Tool ตามที่ต้องการ โดยใช้การไล่สีแบบ Linear Gradient ใช้เครื่องมือ Selection Tool คลิกเลือกเส้นทั้งหมดแล้วกด Delete

เพื่อลบเส้นในภาพออก ตกแต่งภาพโดยใช้ Line Tool จะได้ภาพที่เสร็จสมบูรณ์สามารถนำไปใช้งานได้ ดังภาพที่ 6



ภาพที่ 6: ภาพที่สำเร็จแล้วสามารถนำไปใช้งานได้

3.2.4 ขั้นตอนการทำแอนิเมชันภาพเคลื่อนไหวและใส่เสียงประกอบ

เมื่อได้ออกแบบโครงร่างระบบ และวาดภาพประกอบในส่วนต่างๆ ของโปรแกรมเรียบร้อยแล้ว ขั้นตอนต่อไปคือการทำแอนิเมชันใส่ภาพเคลื่อนไหว การใส่เสียงประกอบ และการเขียนโปรแกรม โดยใช้โปรแกรม Adobe Flash CS6 และ ActionScript เป็นเครื่องมือ

3.2.5 ขั้นตอนการนำโปรแกรมที่เสร็จเรียบร้อยแล้ว ไปติดตั้งลงแท็บเล็ต

เมื่อได้ทำการพัฒนาโปรแกรมเสร็จเรียบร้อยแล้ว ขั้นตอนต่อไปคือทดลองนำโปรแกรมที่เสร็จเรียบร้อยแล้ว ไปใช้จริงบนแท็บเล็ต

3.2.6 ทดสอบการใช้โปรแกรม โดยนำไปให้ผู้ใช้งานทดลองใช้งานจริง

ทดสอบการใช้งานโปรแกรม โดยนำไปให้ผู้ใช้งานทดลองใช้จริง ซึ่งรายละเอียดผลการทดลองใช้โปรแกรมอยู่ในหัวข้อ 4.3 การประเมินผล และ 4.4 การอภิปรายผล

4. ผลการดำเนินงาน

4.1 ผลการพัฒนาโปรแกรม

ผลการพัฒนาโปรแกรม โดยมีตัวอย่างหน้าจอบางส่วนดังต่อไปนี้



ภาพที่ 7: หน้าแรกของโปรแกรม



ภาพที่ 8: หน้าเมนูหลัก

โปรแกรมแบ่งออกเป็น 4 เมนูหลัก เมนูแรก คือ ฟังนิทาน เมื่อคลิกเข้าไปจะเป็นการฟังนิทาน สามารถเลือกฟังเป็นภาษาอังกฤษหรือภาษาไทยก็ได้



ภาพที่ 9: นิทานในรูปแบบภาษาอังกฤษ

เมนูที่สอง คือ เมนูคำศัพท์ ซึ่งจะแบ่งเป็นหมวดคำศัพท์ย่อยอีก 7 หมวด



ภาพที่ 10: เมนูคำศัพท์

เมนูที่สาม คือ เมนูแบบฝึกหัด ซึ่งจะแบ่งเป็น 7 หมวดย่อยเหมือนเมนูคำศัพท์ ในแต่ละหมวดย่อย จะมีแบบฝึกหัดหมวดละ 10 ข้อ เมนูสุดท้าย คือ เมนูเกมสุดสนุก ชื่อเกม ทายชิโครเอ่ย ? เป็นเกมเปิดแผ่นป้าย เพื่อทายชื่อตัวละครในนิทาน ถ้าไม่เปิดแผ่นป้ายเลย แต่ตอบถูกจะได้คะแนนเต็ม ถ้าเปิดแผ่นป้ายคะแนนจะลดลงเรื่อยๆ



ภาพที่ 11: แบบฝึกหัด



ภาพที่ 12: เกมทายชิโครเอ่ยและหน้าสรุปคะแนนที่ได้

4.2 ปัญหาที่พบจากการพัฒนาโปรแกรม

ในระหว่างการศึกษาและพัฒนาโปรแกรม คณะผู้วิจัยได้ประสบปัญหาต่างๆ ดังนี้

1. มีความยากในการวาดฉาก ตัวละคร การสร้างภาพเคลื่อนไหว และการใส่เสียงประกอบให้ตรงกับที่ต้องการ
2. รูปแบบตัวอักษร ไม่ตรงกับที่กำหนดจัดให้เป็นระเบียบยาก
3. ขนาดหน้าจอของระบบแอนดรอยด์ ไม่ตรงกับที่ต้องการ

4.3 การประเมินผล

เมื่อพัฒนาโปรแกรมเสร็จแล้ว คณะผู้วิจัยได้นำโปรแกรมให้ผู้ที่เกี่ยวข้องทดลองใช้งาน เพื่อประเมินผลความพึงพอใจของโปรแกรม

โดยแบ่งระดับการให้คะแนนจากน้อยที่สุด ไปหามากที่สุด ระดับ 1 ถึง 5 (โดย 1 คือพึงพอใจน้อยที่สุด ส่วน 5 คือ พึงพอใจมากที่สุด) จากนั้นให้ผู้ใช้งาน มาทดลองใช้งานโปรแกรมแล้วทำแบบสอบถามเพื่อประเมินผล ซึ่งผู้ทดลองใช้งานได้แก่เด็กในระดับประถมศึกษาตอนปลาย และระดับมัธยมศึกษาตอนต้นจำนวน 50 คน และอาจารย์ที่สอนในระดับประถมศึกษาจำนวน 2 คน

4.4 การอภิปรายผล

ผลการศึกษาจากการตอบแบบสอบถาม โดยการให้ผู้ใช้งานจำนวน 52 คน ทำแบบสอบถามและประเมินผลการใช้งานพบว่า ผลการประเมินความพึงพอใจของโปรแกรมได้ค่าคะแนนเฉลี่ยคิดเป็น 4.08 จากผู้ทดลองใช้งานจำนวน 52 คน และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.04 ซึ่งถือว่าโปรแกรมนี้ มีความพึงพอใจอยู่ในระดับดี สามารถดูรายละเอียดผลการประเมิน ได้จากตารางที่ 1

ตารางที่ 1: การประเมิน โปรแกรม

หัวข้อที่ประเมิน	การวิเคราะห์ข้อมูล		
	ค่าเฉลี่ย	SD.	ความพึงพอใจ
1. ภาพประกอบมีสีสันสวยงาม	3.96	0.62	ดี
2.เสียงประกอบช่วยให้เกิดความน่าสนใจในการใช้โปรแกรม	4.00	0.68	ดี
3.ภาพเคลื่อนไหว animation ช่วยให้เกิดความน่าสนใจในการใช้โปรแกรม	4.29	0.63	ดี
4.เมนูและการออกแบบมีความเหมาะสม	3.94	0.63	ดี
5.ความสนุกสนานเพลิดเพลินจากการใช้งานโปรแกรม	3.96	0.65	ดี
6.ประโยชน์ที่ได้รับจากการใช้งานโปรแกรม	4.10	0.63	ดี
7.โดยภาพรวม มีความพึงพอใจโปรแกรม อยู่ในระดับใด	4.33	0.54	ดี
เฉลี่ย	4.08	0.04	ดี

5. สรุป

หลังจากที่ได้ศึกษา ออกแบบและพัฒนาสื่อการเรียนรู้ คำศัพท์ภาษาอังกฤษบนแท็บเล็ต แล้วทำการทดสอบและประเมินผลการใช้งานกับผู้ใช้จำนวน 52 คน โดยแบ่งเป็นการทดสอบจากอาจารย์ที่สอนในระดับประถมศึกษาจำนวน 2 คน และเด็กในระดับประถมศึกษาตอนปลาย และระดับมัธยมศึกษาตอนต้นจำนวน 50 คน พบว่าสื่อการเรียนรู้ดังกล่าวเป็นประโยชน์ต่อการเรียนรู้ภาษาอังกฤษ โดยเฉพาะอย่างยิ่ง ช่วยเสริมสร้างทักษะภาษาอังกฤษให้กับผู้ใช้ได้เป็นอย่างดี เนื่องจากสื่อการเรียนรู้มีภาพประกอบสีสันสวยงาม รวมทั้งมีเสียงประกอบในระหว่างการศึกษา ทำให้ผู้ใช้มีความสนใจ และได้รับความสนุกสนานเพลิดเพลินไปกับสื่อการเรียนรู้ ไม่ว่าจะเป็นส่วนของนิทาน คำศัพท์ แบบฝึกหัด หรือเกม ซึ่งมีสีสันสวยงาม รวมทั้งมีเสียงประกอบในระหว่าง

การศึกษาทำให้ผู้ใช้มีความสนใจ และได้รับความสนุกสนานเพลิดเพลินไปกับสื่อการเรียนรู้

ข้อเสนอแนะ

จากการนำไปทดลองใช้งาน มีข้อเสนอแนะจากผู้ใช้งานเพิ่มเติม สามารถสรุปได้ดังนี้

1. ควรเพิ่มการเก็บข้อมูล ผลคะแนนของการทำแบบฝึกหัด
2. ควรเพิ่มความสามารถใช้งานแบบออนไลน์ได้
3. ควรเพิ่มนิทานเรื่องอื่นๆและเพิ่มจำนวนคำศัพท์ให้มากขึ้น

เอกสารอ้างอิง

- [1] Wall Street English, “ความสามารถในการใช้ภาษาอังกฤษของคนไทย,” [Online] 2014. Availablefrom: <http://www.wallstreetenglish.in.th/wall-street-english/aec> [2014, March 11].
- [2] จักรชัย โสอินทร์, “Android App Development ฉบับสมบูรณ์,” นนทบุรี : บริษัทไอดีซีพีริเมียร์ จำกัด , 2555.
- [3] ณัฐภูมิ โอภาภิบาล, “คู่มือใช้งาน Android Tablet ฉบับสมบูรณ์,” นนทบุรี : บริษัทไอดีซีพีริเมียร์ จำกัด , 2555.
- [4] Admin, “แท็บเล็ตคืออะไร,” [Online] 2013. Availablefrom: <http://www.tabletd.com/articles/289> [2013,February 4].
- [5] กำพล ธิลาภรณ์, “Flash ActionScript,” กรุงเทพฯ : สำนักพิมพ์ไปรษวิชั่น, 2551.
- [6] อิศเรศ ภาชนะกาญจน์, “Flash CSS basic,” นนทบุรี : บริษัทไอดีซีพีริเมียร์ จำกัด, 2554.
- [7] คุณเดือน กลั่นภู่วัด, “หนูน้อยหมวกแดง,” กรุงเทพฯ : สำนักพิมพ์เอ็มไอเอส, 2553.
- [8] พวงทอง ไกรพิบูลย์, “เด็ก,” [Online] 2014. Availablefrom: <http://haamor.com/th/เด็ก> [2014, March 11].
- [9] สุธมนทา วงษ์สวัสดิ์, “การใช้นิทานเป็นสื่อในการเรียนรู้เพื่อพัฒนาความสามารถด้านการฟังภาษาอังกฤษของนักเรียนชั้นประถมศึกษาปีที่ 5 โรงเรียนประถมนันทรี กรุงเทพมหานคร,” สารนิพนธ์ศิลปศาสตรมหาบัณฑิต มหาวิทยาลัยศรีนครินทรวิโรฒ, 2551.
- [10] เกรียงศักดิ์ สขะนันท์ และวัฒนา พัดเกตุ, “อายุกับการเรียนภาษาอังกฤษและปัจจัยที่มีอิทธิพลต่อความพอใจหรือไม่พอใจในการเรียนภาษาอังกฤษ,” วารสารมนุษยศาสตร์ มหาวิทยาลัยนครสวรรค์, ปีที่ 2 ฉบับที่ 2 กรกฎาคม – ธันวาคม, 2548.

**การประยุกต์ใช้สื่ออีดีแอลทีวีกลุ่มสาระการเรียนรู้
วิทยาศาสตร์สู่การพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนแท็บเล็ต
The Applications of eDLTV Media on Science Learning Strand through the
Application Development for Learning on the Tablet**

อนุกุล บุตรพรหม¹ (Anukul Butprom) , วรปภา อารีราษฎร์² (Worapapha Arreerard)

¹นักศึกษา สาขาคอมพิวเตอร์ศึกษา คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏมหาสารคาม

²อาจารย์ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏมหาสารคาม

Anukul.but@hotmail.com¹, dr.Worapapha@hotmail.com²

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ 1) พัฒนาคู่มืออบรมการประยุกต์ใช้สื่ออีดีแอลทีวีกลุ่มสาระการเรียนรู้วิทยาศาสตร์สู่การพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนแท็บเล็ต 2) เปรียบเทียบผลการเรียนรู้ก่อนอบรมและหลังอบรมของผู้เข้ารับการอบรมด้วยคู่มืออบรมที่พัฒนาขึ้น 3) ศึกษาคุณภาพแอปพลิเคชันเพื่อการเรียนรู้บนแท็บเล็ตของผู้เข้ารับการอบรมด้วยคู่มืออบรมที่พัฒนาขึ้น และ 4) ศึกษาความพึงพอใจของผู้เข้ารับการอบรมที่มีต่อคู่มืออบรมที่พัฒนาขึ้น กลุ่มตัวอย่าง คือ นักศึกษาระดับปริญญาตรีวิชาเอกเคมี ภาคเรียนที่ 1 ปีการศึกษา 2556 คณะครุศาสตร์ มหาวิทยาลัยราชภัฏมหาสารคาม จำนวน 30 คน เครื่องมือที่ใช้ในการวิจัย ได้แก่ คู่มืออบรม แบบทดสอบวัดผลการเรียนรู้ แบบประเมินคุณภาพแอปพลิเคชันเพื่อการเรียนรู้ และแบบประเมินความพึงพอใจที่มีต่อคู่มืออบรม สถิติที่ใช้วิเคราะห์ข้อมูล คือ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐานและสถิติ t-test แบบ (Dependent Sample) ผลการวิจัยพบว่า 1) คู่มืออบรม มีคุณภาพอยู่ในระดับมากที่สุด ($\bar{X}=4.57$, $S.D.=0.50$), 2) ผู้เข้าอบรมมีผลการเรียนรู้หลังการอบรมสูงกว่าก่อนการอบรมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05, 3) ผลการศึกษาคุณภาพแอปพลิเคชันเพื่อการเรียนรู้บนแท็บเล็ตของผู้เข้ารับการอบรมมีคุณภาพ อยู่ในระดับมากถึงมากที่สุด ($\bar{X}=4.00-4.67$, $S.D.=0.39-0.79$) และ 4) ผู้เข้ารับการอบรมมีความพึงพอใจต่อคู่มืออบรมอยู่ในระดับพึงพอใจมากที่สุด ($\bar{X}=4.50$, $S.D.=0.54$)

คำสำคัญ: สื่ออีดีแอลทีวี แอปพลิเคชัน แท็บเล็ต

Abstract

The objectives of this research were to: 1) develop the quality of training course on the applications of eDLTV media on science learning strand through the application development for learning on the tablet, 2) study the result of the students' learning before and after training in the developed course, 3) study the quality of application, and 4) survey the students' satisfaction through the training course. Subjects used in the research consisted of 30 undergraduate students who were majoring in chemistry in the first semester of the 2013 academic year, the Faculty of Education, Rajabhat Maha Sarakham University. The instruments used in the research were a training guide, learning test paper, evaluating an application form, and satisfaction questionnaire. The statistics used for analyzing data were mean, standard deviation and t-test (dependent sample).

The results of the research were as follows: 1) the training course was rated at the highest level ($\bar{X}=4.57$, $S.D.=0.50$), 2) the result of the students' learning after training was significantly higher than before training at the .05 level, 3) regarding the quality of application, the finding revealed that the score was appropriate right up to the most ($\bar{X}=4.00-4.67$, and $S.D.=0.39-0.79$), and 4) the overall students' satisfaction towards the training course was rated at the highest level ($\bar{X}=4.50$, $S.D.=0.54$).

Keyword: eDLTV, Application, Tablet

1. บทนำ

นโยบายการศึกษาของรัฐบาล โดยคณะรัฐมนตรี ได้แถลงนโยบายในส่วนที่เกี่ยวข้องด้านการศึกษาต่อรัฐสภา ในส่วนของนโยบายเร่งด่วนของรัฐบาล ได้แก่ การส่งเสริมให้นักเรียนทุกระดับชั้นได้ใช้อุปกรณ์คอมพิวเตอร์แท็บเล็ตเพื่อการศึกษา [1] การจัดการศึกษาโดยภาครัฐที่กล่าวในเบื้องต้นนั้น “แท็บเล็ตเพื่อการศึกษา (Tablet for Education)” จึงกลายเป็นเครื่องมือด้านสื่อเทคโนโลยีเพื่อการศึกษาที่สำคัญและมีอิทธิพลค่อนข้างมากต่อการพัฒนาการจัดการศึกษาไทยในปัจจุบัน ซึ่งแนวนโยบายของรัฐบาลมุ่งเน้นที่จะใช้สื่อแท็บเล็ตให้ผู้เรียนทุกคนได้เรียนรู้ตามศักยภาพและความพร้อมที่มีอยู่ โดยที่นโยบายการปฏิบัติกับนักเรียนช่วงแรกตามโครงการ One Tablet PC Per Child จะมุ่งเน้นไปที่นักเรียนชั้นประถมศึกษาปีที่ 1 เป็นกลุ่มเป้าหมายนำร่องที่สำคัญของการนำสื่อแท็บเล็ตสู่การพัฒนาการเรียนรู้ในครั้งนี้ [2]

สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐานเป็นหน่วยงานที่นำนโยบายของรัฐบาลลงสู่ปฏิบัติ ในปี 2555 ดำเนินการพัฒนาศักยภาพของผู้ที่เกี่ยวข้อง ได้แก่ ผู้บริหารเขตพื้นที่ ผู้บริหารโรงเรียน ศึกษานิเทศก์ ครู และผู้ดูแลระบบ เพื่อให้สามารถใช้คอมพิวเตอร์แท็บเล็ตได้อย่างมีประสิทธิภาพ และในปี 2556 ได้ดำเนินโครงการ “ประกวดสร้างสื่อการเรียนรู้...สู่แท็บเล็ต” เพื่อเป็นการส่งเสริมให้มีสื่อเทคโนโลยีที่มีคุณภาพ รวมทั้งเป็นการสนับสนุนและกระตุ้นให้ครู บุคลากรทางการศึกษา และประชาชนทั่วไป มีการพัฒนาสื่อเพื่อการเรียนรู้สำหรับนักเรียนรูปแบบแอปพลิเคชันบนระบบปฏิบัติการแอนดรอยด์ที่มีคุณภาพ และปริมาณเพียงพอต่อการใช้งาน [3] สทฐ. ได้ดำเนินการพัฒนาและจัดเตรียมเครื่องมือสำหรับสร้างสื่อแอปพลิเคชัน เพื่อเป็นเครื่องมือในการพัฒนาแอปพลิเคชันได้อย่างมีประสิทธิภาพ ในขณะเดียวกัน ได้มีการจัดอบรมเชิงปฏิบัติการ ให้กับครูทั่วประเทศ จำนวน 1,000 คน ให้มีความรู้และทักษะในการพัฒนาสื่อแอปพลิเคชัน และได้ดำเนินการจัดโครงการอบรมการสร้างสื่อการเรียนรู้แอปพลิเคชันให้กับครูทั่วประเทศ ใน 5 ภูมิภาค จำนวน 5 รุ่น โดยจัดอบรมรุ่นละ 200 คน รวมทั้งสิ้น 1,000 คน และเนื่องจากในการจัดการอบรมครั้งนี้มีครู และบุคลากรทางการศึกษาให้ความสนใจสมัครเข้าร่วมเป็นจำนวนมาก เดิมทีจะเปิดการอบรม จึงทำให้ครู และบุคลากรทางการศึกษาที่สนใจสมัครเข้าร่วมอบรมพลาดโอกาสนี้ไป

มหาวิทยาลัยราชภัฏมหาสารคาม เป็นสถาบันการศึกษาเพื่อพัฒนาชุมชนและท้องถิ่น ส่งเสริมการพัฒนาชุมชนเพื่อการเรียนรู้ด้วยสื่อไอซีที โดยคณะเทคโนโลยีสารสนเทศได้ดำเนินการร่วมมือกับเครือข่ายทั้งหน่วยงานภาครัฐและเอกชน ดำเนินงานโครงการส่งเสริมการใช้สื่ออีดีแอลทีวี หรือ eDLTV ซึ่งเป็นสื่อภายใต้โครงการจัดทำเนื้อหาบนระบบอีดีแอลทีวี eDL-Square ภายใต้โครงการเทคโนโลยีสารสนเทศตามพระราชดำริ สมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี เพื่อให้ครูและบุคลากรทางการศึกษาได้ใช้เป็นสื่อเพื่อการเรียนการสอน [4]

มหาวิทยาลัยราชภัฏมหาสารคามร่วมลงนามบันทึกข้อตกลงความร่วมมือร่วมกับสำนักงานเขตพื้นที่การศึกษามหาสารคาม เขต 1 ได้ดำเนินการอบรม เพื่อพัฒนาครูและบุคลากรทางการศึกษา ในการสร้างแอปพลิเคชันการเรียนรู้ เพื่อส่งเสริมพัฒนาครูในการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนการสอนและการประยุกต์ใช้สื่ออีดีแอลทีวีเพื่อการเรียนรู้ [5] จากการจัดอบรมที่ผ่านมา พบว่า ยังไม่ครอบคลุม เนื่องจาก ยังมีครูและบุคลากรทางการศึกษาอีกจำนวนมากที่สนใจสมัครเข้าร่วมการอบรม รวมทั้งนักศึกษา ควรได้รับโอกาสในการพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนแท็บเล็ต เพื่อให้สามารถพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนแท็บเล็ต ที่เป็นสื่อในการเรียนการสอนให้กับนักเรียนได้อย่างมีคุณภาพ ดังนั้นผู้วิจัยได้เล็งเห็นความสำคัญของปัญหาในจุดนี้ จึงได้ดำเนินการจัดโครงการอบรม ขึ้นมาเพื่อสนองความต้องการสำหรับนักศึกษาที่สนใจสมัครเข้าร่วมการอบรมในครั้งนี้

จากหลักการและเหตุผลดังกล่าวข้างต้น ผู้วิจัย จึงเห็นว่ามีออบรมการประยุกต์ใช้สื่ออีดีแอลทีวีกลุ่มสาระการเรียนรู้วิทยาศาสตร์สู่การพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนแท็บเล็ต โดยการศึกษาเนื้อหาสาระการเรียนรู้จากสื่ออีดีแอลทีวี เพื่อให้นักศึกษาสามารถสร้างแอปพลิเคชันเพื่อการเรียนรู้บนแท็บเล็ต เป็นสื่อในการเรียนการสอน ให้กับนักเรียนได้อย่าง มีคุณภาพ ส่งผลให้นักเรียนมีสื่อที่สามารถใช้ในการเรียนรู้บนแท็บเล็ตเพิ่มมากขึ้น

2. วัตถุประสงค์

2.1 เพื่อพัฒนาคู่มืออบรมการประยุกต์ใช้สื่ออีดีแอลทีวีกลุ่มสาระการเรียนรู้วิทยาศาสตร์สู่การพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนแท็บเล็ต ให้มีคุณภาพ

2.2 เพื่อเปรียบเทียบผลการเรียนรู้ก่อนอบรมและหลังอบรมของผู้เข้ารับการอบรมด้วยคู่มืออบรมที่พัฒนาขึ้น

2.3 เพื่อศึกษาคุณภาพแอปพลิเคชันของผู้เข้ารับการอบรมด้วยคู่มืออบรมที่พัฒนาขึ้น

2.4 เพื่อศึกษาความพึงพอใจของผู้เข้ารับการอบรมที่มีต่อคู่มืออบรมที่พัฒนาขึ้น

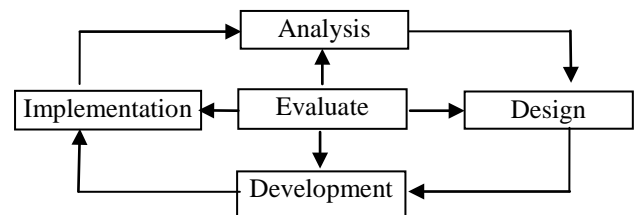
3. วรรณกรรมที่เกี่ยวข้อง

3.1 **สื่ออีดีแอลทีวี** เป็นสื่อที่มูลนิธิการศึกษาทางไกลผ่านดาวเทียมและโครงการเทคโนโลยีสารสนเทศตามพระราชดำริสมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี จัดทำขึ้นเพื่อร่วมเทิดพระเกียรติพระบาทสมเด็จพระเจ้าอยู่หัว เนื่องในโอกาสสมหามงคลเฉลิมพระชนมพรรษา 80 พรรษา 5 ธันวาคม 2550 โดยบรรจุเนื้อหาสาระบทเรียนจากการสอนในระดับมัธยมศึกษาของโรงเรียนไกลกังวล จำนวน 6 กลุ่มสาระการเรียนรู้ คือ ภาษาไทย วิทยาศาสตร์ คณิตศาสตร์ สังคมศึกษา สุขศึกษา และภาษาอังกฤษ โดยจัดเป็นสื่อวีดิทัศน์การสอน ประกอบด้วยสไลด์ประกอบการสอน ใบงาน ใบความรู้ และแบบทดสอบ รวบรวมและจัดเก็บด้วยระบบบริหารจัดการเรียนรู้อีดีแอลสแควร์ สามารถจัดการเรียนการสอนโดยใช้การศึกษาทางไกลผ่านดาวเทียมในรูปแบบสื่ออิเล็กทรอนิกส์เพื่อการเรียนรู้แบบไม่ผ่านเครือข่าย หรืออีเลิร์นนิ่งแบบออนไลน์ หรือเพื่อให้อาจจัดการเรียนการสอนได้ทั้งแบบออนไลน์ ผ่านระบบอินเทอร์เน็ต ให้แก่ครู นักเรียน และผู้ที่สนใจทั่วไปได้ใช้ประโยชน์ในการเรียนการสอนหรือศึกษาเพิ่มเติม [6]

3.2 **แอปพลิเคชันเพื่อการเรียนรู้** นโยบายด้านการศึกษารัฐในการจัดสรรเครื่องคอมพิวเตอร์แท็บเล็ตให้กับนักเรียนชั้นประถมศึกษาปีที่ 1 ในภาคเรียนที่ 1 ในปีการศึกษา 2555 โดยมีวัตถุประสงค์เพื่อส่งเสริมการใช้เทคโนโลยีเป็นเครื่องมือยกระดับคุณภาพและกระจายโอกาสทางการศึกษาให้กับนักเรียนทั่วประเทศ โดย สพฐ.รับผิดชอบด้านการพัฒนาสื่อเพื่อการเรียนรู้ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 ดำเนินการ จัดโครงการประกวดสร้างสื่อการเรียนรู้สู่แท็บเล็ตเพื่อส่งเสริมให้มีสื่อที่มีคุณภาพ รวมทั้งเป็นการสนับสนุนและกระตุ้นให้ครู บุคลากรทางการศึกษา และ

ประชาชนทั่วไป มีการพัฒนาสื่อเพื่อการเรียนรู้สำหรับนักเรียนรูปแบบแอปพลิเคชันบนระบบปฏิบัติการแอนดรอยด์ ที่มีคุณภาพและปริมาณเพียงพอต่อการใช้งาน สรุปได้ว่า การจัดการศึกษาไทยในปัจจุบันตามแนวนโยบายของรัฐบาลมุ่งเน้นที่จะใช้สื่อแท็บเล็ตให้ผู้เรียนทุกคนได้เรียนรู้ตามศักยภาพตามโครงการ One Tablet PC Per Child โดยนำเอาเทคโนโลยี และสื่อสารการศึกษามาประยุกต์ ใช้กับการเรียนรู้ของนักเรียนในรูปแบบใหม่โดยการใช้แท็บเล็ต (Tablet) เป็นเครื่องมือในการเข้าถึงแหล่งเรียนรู้และแสวงหาองค์ความรู้ในรูปแบบต่างๆ ที่มีอยู่ในรูปแบบทั้ง ออฟไลน์และออนไลน์ ทำให้ผู้เรียนมีโอกาสศึกษาหาความรู้ ฝึกปฏิบัติ และสร้างองค์ความรู้ต่างๆ ได้ด้วยตัวเอง

3.3 **การพัฒนาคู่มืออบรม** รูปแบบ ADDIE เป็นรูปแบบที่ได้รับการยอมรับอย่างกว้างขวางในการนำมาใช้ในการพัฒนาหลักสูตรโดยรูดเคอริค ซิมส์(Roderic Sime)มหาวิทยาลัยซิดนีย์ (University of Technology Sydney) ได้นำรูปแบบ ADDIE Model มาปรับปรุงขั้นตอนที่ครอบคลุมสาระสำคัญในการออกหลักสูตรทั้งหมด [7] แสดงดังแผนภาพที่ 1



ภาพที่ 1: ขั้นตอนการพัฒนาหลักสูตรตามรูปแบบ ADDIE Model

จากแผนภาพที่ 1 จะเห็นว่า การพัฒนาหลักสูตรตามรูปแบบ ADDIE Model ประกอบด้วยขั้นตอนทั้งหมด 5 ขั้นตอน ได้แก่ ขั้นตอนการวิเคราะห์ (Analysis) เป็นขั้นวางแผนหรือเตรียมการสื่อต่างๆ ที่จำเป็นต่อการพัฒนาหลักสูตร ขั้นการออกแบบ (Design) เป็นขั้นที่นำข้อมูลต่างๆ ที่ได้วิเคราะห์ไว้เป็นข้อมูลพื้นฐานในการออกแบบ ขั้นการพัฒนา (Development) เป็นขั้นที่นำสิ่งต่างๆ ที่ได้ทำการออกแบบไว้ มาพัฒนา เมื่อดำเนินการพัฒนาหลักสูตรแล้ว ผู้ออกแบบจะต้องนำหลักสูตรนั้นไปทดสอบเพื่อตรวจสอบหาความผิดพลาดและเพื่อความสมบูรณ์ของแต่ละโมดูล ขั้นการทดลองใช้ (Implementation) เป็นขั้นที่นำหลักสูตรที่มีองค์สมบูรณ์มาทดลองใช้เพื่อหาประสิทธิภาพของหลักสูตร และขั้นประเมินผล (Evaluate) เป็นการนำผลการทดลองที่ได้มาสรุปผล

ซึ่งกระบวนการทั้ง 5 ขั้นทำให้ได้หลักสูตรครบถ้วนสมบูรณ์เหมาะสมกับการอบรมได้เป็นอย่างดี ผู้วิจัยจึงนำมาเป็นแนวทางในการพัฒนาคู่มืออบรมในการวิจัยในครั้งนี้

3.4 งานวิจัยที่เกี่ยวข้อง

นางเยาว์ ประธิปอาราม [8] ได้ทำวิจัยเรื่อง การพัฒนาหลักสูตรอบรมการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ เรื่อง การตกแต่งรูปภาพโดยใช้ Ps Express ด้วย Ipad2 สำหรับชั้นมัธยมศึกษาปีที่ 4 ผลการวิจัยสรุปได้ดังนี้ 1) การพัฒนาหลักสูตรอบรม ผลการวิเคราะห์ความคิดเห็นของผู้เชี่ยวชาญที่มีต่อการพัฒนาหลักสูตรอบรมโดยภาพรวมอยู่ในระดับเหมาะสมมาก 2) ผลการเรียนรู้ของผู้เข้ารับการอบรมด้วยหลักสูตรที่พัฒนาขึ้นสรุปได้ว่า ผู้เข้ารับการอบรมมีความรู้ความเข้าใจในการอบรม โดยมีคะแนนหลังอบรมสูงกว่าก่อนอบรมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และ 3) ความพึงพอใจของผู้เรียนที่ได้รับการอบรมด้วยหลักสูตรอบรมที่พัฒนาขึ้น พบว่า นักเรียนมีความพึงพอใจอยู่ในระดับมาก

อรณัน เฝื่อนผึ้ง [9] ได้ทำวิจัยเรื่อง การพัฒนาหลักสูตรอบรมการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ เรื่อง การทำภาพสติกเกอร์ด้วย Ipad2 สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 1 ผลการวิจัยสรุปได้ดังนี้ 1) การพัฒนาคุณภาพหลักสูตรอบรม พบว่า ผู้เชี่ยวชาญมีความคิดเห็นโดยรวมต่อหลักสูตรอบรมอยู่ในระดับเหมาะสมมาก 2) ผลการเปรียบเทียบผลการเรียนรู้ก่อนอบรมและหลังอบรมของผู้เข้ารับการอบรมด้วยหลักสูตรอบรมที่พัฒนาขึ้นพบว่า ผู้เข้ารับการอบรมมีความรู้ความเข้าใจในการอบรม โดยมีคะแนนหลังอบรมสูงกว่าก่อนอบรมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และ 3) ความพึงพอใจของผู้เข้ารับการอบรมด้วยหลักสูตรอบรมที่พัฒนาขึ้น เฉลี่ยโดยรวมอยู่ในระดับมากที่สุด

นำชัย โบราณมูล [10] ได้ศึกษาวิจัยเรื่อง การพัฒนาหลักสูตรอบรมการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ เรื่อง การสร้างสเปรดชีตด้วยแอปพลิเคชัน Number สำหรับนักศึกษา ผลการวิจัยสรุปได้ดังนี้ 1) การพัฒนาหลักสูตรอบรม ผลการวิเคราะห์ความคิดเห็นของผู้เชี่ยวชาญต่อการพัฒนาหลักสูตรอบรมในภาพรวมอยู่ในระดับเหมาะสมมาก 2) ผลการเปรียบเทียบผลการเรียนรู้ก่อนอบรมและหลังอบรมของผู้เข้ารับการอบรมด้วยหลักสูตรอบรมที่พัฒนาขึ้น พบว่า ผู้เข้ารับการอบรมมี

ความรู้ความเข้าใจในการอบรม โดยมีคะแนนหลังอบรมสูงกว่าก่อนอบรมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และ 3) ความพึงพอใจของผู้เข้ารับการอบรมที่มีต่อหลักสูตรอบรม พบว่า ผู้เรียนมีความพึงพอใจต่อหลักสูตรอบรมที่พัฒนาขึ้น เฉลี่ยโดยรวมส่วนใหญ่อยู่ในระดับพึงพอใจมาก

นิโคลัส จอร์จฟัสกี และคณะ. (2005 : 95-102) [11] ได้ทำการศึกษาเรื่อง เครื่องมือสำหรับการเรียนการสอนแคลคูลัส พบว่า การรับรู้ของนักเรียนจากเครื่อง Tablet PC เป็นเครื่องมือในการเรียนการสอนสำหรับการสอนแคลคูลัส สำรวจรายการสิบสามรายการ ซึ่งวัตถุประสงค์ของการสำรวจครั้งนี้ คือ การเก็บรวบรวมข้อมูลเกี่ยวกับการรับรู้ของนักเรียนเกี่ยวกับการใช้งานของเครื่อง Tablet PC เป็นเครื่องมือในการเรียนการสอนได้อย่างมีประสิทธิภาพ และมีประสิทธิภาพครอบคลุมวัตถุประสงค์การเรียนการสอนในชั้นเรียน ผลการชี้ให้เห็นว่านักเรียนรับรู้ว่าการใช้เครื่อง Tablet PC ทำให้พวกเขามีความสนใจในชั้นเรียน ช่วยให้เข้าใจเนื้อหาในเรื่องที่เรียน ครอบคลุมเนื้อหาการเรียน ทำให้เกิดประสิทธิภาพในการเรียน

4. ระเบียบวิธีวิจัย

การวิจัยเรื่อง การประยุกต์ใช้สื่ออีดีแอลทีวีในกลุ่มสาระการเรียนรู้วิทยาศาสตร์สู่การพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนแท็บเล็ต ผู้วิจัยได้ดำเนินการวิจัยดังนี้

4.1 กลุ่มเป้าหมายการวิจัย คือ นักศึกษาระดับปริญญาตรี วิชาเอกเคมี ภาคเรียนที่ 1 ปีการศึกษา 2556 คณะครุศาสตร์ มหาวิทยาลัยราชภัฏมหาสารคาม จำนวน 30 คน ที่เข้ารับการอบรม

4.2 ขั้นตอนการการวิจัย

ผู้วิจัยได้ดำเนินการตามขั้นตอน ADDIE Model 5 ขั้นตอน ดังนี้

4.2.1 ขั้นการวิเคราะห์ ผู้วิจัยดำเนินการศึกษาเอกสารงานวิจัยที่เกี่ยวข้อง ศึกษาในด้านการพัฒนาแอปพลิเคชัน และการประยุกต์ใช้สื่ออีดีแอลทีวี

4.2.2 ขั้นการออกแบบ ผู้วิจัยดำเนินการออกแบบคู่มืออบรมการประยุกต์ใช้สื่ออีดีแอลทีวี กลุ่มสาระการเรียนรู้วิทยาศาสตร์ สู่การพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนแท็บเล็ต เครื่องมือการประเมินผลต่าง ๆ

4.2.3 ขั้นการพัฒนา ผู้วิจัยดำเนินการพัฒนาคู่มืออบรม และเครื่องมือแบบประเมิน ต่าง ๆ พร้อมทั้งหาคุณภาพคู่มือ อบรม และแบบประเมินผลต่าง ๆ

4.2.4 ขั้นการทดลองใช้ ผู้วิจัยดำเนินการทดลองกับ กลุ่มเป้าหมายตามแบบแผนการทดลอง เก็บรวบรวมข้อมูล ต่าง ๆ ที่เกี่ยวข้อง

4.2.5 ขั้นการประเมินผล ดำเนินการนำข้อมูลที่ได้จาก การทดลองวิเคราะห์หาค่าทางสถิติ และสรุปผลการทดลอง เขียนรายงานผลการวิจัย

4.3 แบบแผนการทดลอง

การวิจัยครั้งนี้ มีแบบแผนการทดลอง ดังตารางที่ 1

ตารางที่ 1: แบบแผนการทดลอง

E	T ₁	X	T ₂
---	----------------	---	----------------

โดยที่ E หมายถึง กลุ่มทดลอง
 T1 หมายถึง ทดสอบก่อนการทดลอง
 T2 หมายถึง ทดสอบหลังการทดลอง
 X หมายถึง จัดอบรมโดยใช้คู่มืออบรมที่ พัฒนาขึ้น

4.4 เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้

4.4.1 คู่มืออบรมการประยุกต์ใช้สื่ออีดีแอลทีวี กลุ่มสาระ การเรียนรู้วิทยาศาสตร์สู่การพัฒนาแอปพลิเคชันเพื่อการเรียนรู้ บนแท็บเล็ต

4.4.2 แบบประเมินคุณภาพคู่มืออบรม

4.4.3 แบบทดสอบวัดผลการเรียนรู้

4.4.4 แบบประเมินคุณภาพแอปพลิเคชันเพื่อการเรียนรู้บน แท็บเล็ต ของผู้รับการอบรม

4.4.5 แบบสอบถามความพึงพอใจของผู้รับการอบรม

5. ผลการวิจัย

การวิจัยครั้งนี้ เป็นการศึกษาการประยุกต์ใช้สื่ออีดีแอลทีวี กลุ่มสาระการเรียนรู้วิทยาศาสตร์สู่การพัฒนาแอปพลิเคชันเพื่อ การเรียนรู้บนแท็บเล็ต ผู้วิจัยได้นำเสนอเป็นลำดับ ดังนี้

5.1 ผลการศึกษาคุณภาพคู่มืออบรม ผู้วิจัยได้พัฒนาคู่มือ อบรมการประยุกต์ใช้สื่ออีดีแอลทีวี กลุ่มสาระการเรียนรู้ วิทยาศาสตร์สู่การพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนแท็บเล็ต

และนำคู่มือนำเสนอผู้เชี่ยวชาญจำนวน 5 ท่าน โดยเป็นอาจารย์ จากมหาวิทยาลัยราชภัฏมหาสารคาม ศึกษาบัณฑิตศึกษาและครูจาก สำนักงานเขตพื้นที่การศึกษามหาสารคาม และเป็นผู้ซึ่งได้ผ่าน การอบรมการเป็นวิทยากรการสร้างแอปพลิเคชัน จาก สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน เพื่อศึกษาความ คิดเห็นผู้เชี่ยวชาญที่มีต่อคู่มือ ผลการวิเคราะห์คุณภาพดังแสดง ในตารางที่ 2

ตารางที่ 2: ผลการวิเคราะห์หาคุณภาพ

รายการ	\bar{X}	S.D.	ความหมาย
1. หลักการและเหตุผล	4.53	0.52	มากที่สุด
2. วัตถุประสงค์ของหลักสูตรอบรม	4.67	0.49	มากที่สุด
3. โครงสร้างของหลักสูตรอบรม	4.55	0.51	มากที่สุด
4. ด้านการจัดการ/กิจกรรมการอบรม	4.60	0.50	มากที่สุด
5. การวัดผลการอบรม	4.50	0.53	มากที่สุด
6. การประเมินผลการอบรม	4.50	0.53	มากที่สุด
ความเหมาะสมเฉลี่ยโดยรวม	4.57	0.50	มากที่สุด

จากตารางที่ 2 พบว่า คิดเห็นของผู้เชี่ยวชาญที่มีต่อ คู่มืออบรมโดยภาพรวมอยู่ในระดับมากที่สุด ($\bar{X} = 4.57, S.D. = 0.50$) เมื่อพิจารณารายด้านพบว่า มีคุณภาพอยู่ในระดับมาก ที่สุดทุกด้าน

5.2 ผลการเปรียบเทียบการเรียนรู้ก่อนอบรมและหลังอบรม ของผู้เข้ารับการอบรม ผู้วิจัยทดสอบการเรียนรู้ของ กลุ่มเป้าหมายโดยทำการทดสอบก่อนอบรมและหลังอบรม และนำคะแนนมาวิเคราะห์การเปรียบเทียบด้วยสถิติ t-test (dependent) ผลการวิเคราะห์ดังแสดงในตารางที่ 3

ตารางที่ 3: ผลการวิเคราะห์เปรียบเทียบผลการเรียนรู้ก่อนอบรมและ หลังอบรม

คะแนนการ ทดสอบ	ผู้เข้ารับการ อบรม	\bar{X}	S.D.	t
ก่อนอบรม	30	12.23	10.59	12.59
หลังอบรม	30	20.43	23.08	

ระดับนัยสำคัญ .05

จากตารางที่ 3 พบว่า ค่า t ที่คำนวณได้ มีค่าเท่ากับ 12.59 ซึ่งมีค่ามากกว่าค่า t (ตาราง $\alpha = .05, df = 29$) มีค่าเท่ากับ 1.699 สรุปได้

ว่า คะแนนความรู้ความเข้าใจเฉลี่ยหลังการอบรมสูงกว่าก่อนอบรม อย่างมีนัยสำคัญทางสถิติ .05

5.3 ผลการศึกษาคุณภาพแอปพลิเคชันของผู้เข้ารับการอบรม ผู้วิจัยนำแอปพลิเคชันของกลุ่มเป้าหมาย ที่พัฒนาขึ้นระหว่างการเข้ารับการอบรม จำนวน 30 ชิ้นงาน นำเสนอโดยผู้เชี่ยวชาญจำนวน 5 ท่าน ซึ่งมีคุณสมบัติตามที่กล่าวในหัวข้อที่ 5.1 เพื่อศึกษาคุณภาพของแอปพลิเคชัน ผลการศึกษาพบว่า แอปพลิเคชันของผู้เข้ารับการอบรม จำนวน 30 คน ที่พัฒนาขึ้น มีคุณภาพ เฉลี่ยโดยรวมอยู่ในระดับมาก ถึง มากที่สุด โดยมีค่าเฉลี่ย ($\bar{X} = 4.00-4.67$, S.D.= 0.39-0.79)

5.4 ผลการศึกษาความพึงพอใจของผู้เข้ารับการอบรมหลังการอบรม หลังจากที่ผู้วิจัยทดลองใช้คู่มือกับกลุ่มเป้าหมายแล้ว ผู้วิจัยศึกษาความพึงพอใจ โดยใช้แบบสอบถามความพึงพอใจของผู้เข้ารับการอบรม ผลการศึกษาแสดงในตารางที่ 4

ตารางที่ 4: ผลการวิเคราะห์ความพึงพอใจของผู้เข้ารับการอบรม

รายการ	\bar{X}	S.D.	ความหมาย
1. ด้านวิทยากร	4.50	0.50	มากที่สุด
2. ด้านกระบวนการอบรม	4.52	0.50	มากที่สุด
3. ด้านความรู้ความเข้าใจ	4.50	0.50	มากที่สุด
4. ด้านการนำไปใช้ประโยชน์	4.62	0.49	มากที่สุด
5. ด้านสถานที่ระยะเวลา	4.37	0.73	มาก
ความพึงพอใจเฉลี่ยโดยรวม	4.50	0.54	มากที่สุด

จากตารางที่ 4 พบว่า ความพึงพอใจของผู้เข้ารับการอบรมพบว่า ผู้เข้ารับการอบรมส่วนใหญ่มีความพึงพอใจต่อการจัดอบรมโดยภาพรวม พบว่า อยู่ในระดับมากที่สุด ($\bar{X} = 4.50$, S.D. = 0.54) เมื่อพิจารณารายด้านทั้ง 5 ด้าน พบว่า ทุกด้านมีความพึงพอใจอยู่ในระดับพึงพอใจมากที่สุด

6. สรุปผลการวิจัย

ผลการวิจัยการอบรมด้วยคู่มืออบรมการประยุกต์ใช้สื่ออีดีแอลทีวี กลุ่มสาระการเรียนรู้วิทยาศาสตร์สู่การพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนแท็บเล็ต จากผลการวิจัยสรุปได้ว่า 1) การพัฒนาคู่มืออบรมที่ผู้วิจัยพัฒนาขึ้น อยู่ในระดับเหมาะสมมากที่สุด 2) ผู้เข้าอบรมมีผลการเรียนรู้หลังการอบรมสูงกว่าก่อนการอบรมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 3)

ผลการพัฒนาแอปพลิเคชันเพื่อการเรียนรู้บนแท็บเล็ตของผู้เข้ารับการอบรม มีคุณภาพอยู่ในระดับมาก ถึง มากที่สุด 4) ผู้เข้ารับการอบรมมีความพึงพอใจต่อคู่มืออบรม อยู่ในระดับพึงพอใจมาก

เอกสารอ้างอิง

- [1] สำนักงานเลขาธิการรัฐมนตรี. “คำแถลงนโยบายของคณะรัฐมนตรีนางสาวยิ่งลักษณ์ ชินวัตร” กรุงเทพฯ : สำนักพิมพ์คณะรัฐมนตรีและราชกิจจานุเบกษา, 2554.
- [2] สุรศักดิ์ ปาเฮ. “แท็บเล็ตเพื่อการศึกษา: โอกาสและความท้าทาย” สืบค้นเมื่อ 6 เมษายน 2556 จาก www.kan1.go.th/tablet-for-education.pdf,2554.
- [3] www.otpcappcon.com. (ออนไลน์) สืบค้นเมื่อ 10 มีนาคม 2556
- [4] สมเจตน์ ภูศรี “รายงานการดำเนินงานเครือข่ายเผยแพร่ถ่ายทอดและพัฒนาสื่อการเรียนการสอนบนระบบ e-Learning (eDL- Square)”. มหาวิทยาลัยราชภัฏมหาสารคาม, 2555.
- [5] มหาวิทยาลัยราชภัฏมหาสารคาม .“รายงานผลการดำเนินงานการส่งเสริมการพัฒนาครู บุคลากรทางการศึกษา และนักเรียนในการใช้เทคโนโลยีสารสนเทศเพื่อยกระดับการเรียนรู้ของชุมชน มหาวิทยาลัยราชภัฏมหาสารคาม”. มหาสารคาม : มหาวิทยาลัยราชภัฏมหาสารคาม ,2555.
- [6] เครือข่ายเผยแพร่และถ่ายทอดการใช้สื่ออีดีแอลทีวีในการเรียนการสอน มหาวิทยาลัยราชภัฏมหาสารคาม .[ออนไลน์] สืบค้นจาก <http://www.itrmu.net/edltv/> เมื่อวันที่ 20 มีนาคม 2556
- [7] พิสุทธิ อาวีราษฎร์ .“การพัฒนาซอฟต์แวร์ทางการศึกษา” มหาสารคาม : มหาวิทยาลัยราชภัฏมหาสารคาม, 2551.
- [8] นางเยาว์ ประธิปอราม.“การพัฒนาหลักสูตรอบรมการใช้คอมพิวเตอร์แท็บเล็ต เพื่อการเรียนรู้ เรื่องการตกแต่งรูปภาพโดยใช้ Ps Express ด้วย Ipad2” มหาวิทยาลัยราชภัฏมหาสารคาม, 2555.
- [9] อรฉัตร เตื่อนพั้ง.“การพัฒนาหลักสูตรอบรมการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ เรื่อง การทำภาพสติกเกอร์ ด้วย Ipad2” มหาวิทยาลัยราชภัฏมหาสารคาม, 2555.
- [10] นำชัย โบราณมูล. “การพัฒนาหลักสูตรอบรมการใช้คอมพิวเตอร์แท็บเล็ตเพื่อการเรียนรู้ เรื่องการสร้างสปรดชีทด้วย Numbers สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3”มหาวิทยาลัยราชภัฏมหาสารคาม, 2555.
- [11] Nicholas Gorgievski and other. “Tablet PC: A Preliminary Report on a Tool for Teaching Calculus”. *The International Journal for Technology in Mathematics Education*, 95-102, 2005.

**การพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนแบบเกม 3 มิติ ในรายวิชาคอมพิวเตอร์
ชั้นประถมศึกษาปีที่ 2 กรณีศึกษาโรงเรียนวัดลาดปลาเค้า จังหวัดนครปฐม
A Development of 3D Game Computer Assisted Instruction
In Computer for Grade 2 Students. Case Study at Watladplakao School**

จรินทร์ อุ่มไกร (Charinthorn Aumgri)¹

สาขาวิชาคอมพิวเตอร์ศึกษา คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครปฐม

¹aor030@hotmail.com.com

บทคัดย่อ

งานวิจัยนี้เป็นการวิจัยเชิงทดลองซึ่งมีวัตถุประสงค์
1) เพื่อนำเสนอแนวทางการพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนแบบเกม 3 มิติในรายวิชาคอมพิวเตอร์ ชั้นประถมศึกษาปีที่ 2 กรณีศึกษาโรงเรียนวัดลาดปลาเค้า 2) เพื่อหาความพึงพอใจของบทเรียนคอมพิวเตอร์ช่วยสอนที่พัฒนาขึ้น 3) เพื่อหาประสิทธิภาพของบทเรียนคอมพิวเตอร์ช่วยสอนแบบเกม 3 มิติ ในภาคเรียนที่ 1/2556 เครื่องมือในการวิจัยประกอบด้วยบทเรียนคอมพิวเตอร์ช่วยสอนแบบเกม 3 มิติในรายวิชาคอมพิวเตอร์ ชั้นประถมศึกษาปีที่ 2 กรณีศึกษาโรงเรียนวัดลาดปลาเค้าที่พัฒนาขึ้นแบบสอบถามความพึงพอใจ และแบบทดสอบก่อนและหลังเรียน โดยกลุ่มตัวอย่างในการวิจัยครั้งนี้คือ นักเรียนชั้นประถมศึกษาปีที่ 2 จำนวน 30 คน โดยใช้วิธีการเลือกสุ่มอย่างง่าย ผลการวิจัยพบว่า 1) ความพึงพอใจของการใช้บทเรียนคอมพิวเตอร์ช่วยสอนแบบเกม 3 มิติ ของนักเรียน มีค่าเฉลี่ย 4.39 อยู่ในระดับดี และค่าเบี่ยงเบนมาตรฐาน 0.62 2) บทเรียนคอมพิวเตอร์ช่วยสอนแบบเกม 3 มิติในรายวิชาคอมพิวเตอร์ มีค่าเท่ากับ 85.21/80.06 สรุปได้ว่าบทเรียนคอมพิวเตอร์ช่วยสอนแบบเกม 3 มิติ สามารถนำไปประยุกต์ในรายวิชาคอมพิวเตอร์สำหรับนักเรียนชั้นประถมศึกษาปีที่ 2 ได้
คำสำคัญ: บทเรียนคอมพิวเตอร์ช่วยสอน เกม 3 มิติ คอมพิวเตอร์

Abstract

This research was an experimental research. The purposes of this study were 1) to develop CAI using 3D Game (3DGame-CAI) on Computer for grade 2 students 2) to study satisfaction of the developed 3DGame-CAI and 3) to study the students' achievement after learning

with the developed 3DGame-CAI The sample, selected by simple random sampling, were 3 teachers and 30 students who studied in grade 2 of academic year 2013 at Watladpalkao School. The research tools consisted of the developed 3DGame-CAI on computer, pretest posttest and satisfaction .

The results showed that 1) the evaluation of the satisfaction is a good level and the standard deviation is 4.39, and the average 0.62, 2) the efficiency of 3DGame-CAI was at 85.21/80.06 3) the academic achievement of students learned from the developed 3DGame-CAI was higher than the pretest with the statistical significant of .01. Then we can use this 3DGame-CAI on computer to the target group.

Keyword: CAI, 3DGame, Computer

1. บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ในการเรียนการสอนของนักเรียนโดยใช้เทคโนโลยีกระตุ้นให้เกิดการดึงดูดความสนใจในเรื่องของการเรียนและการจดจำให้ผู้เรียนเกิดทักษะการเรียนรู้ สนใจการเรียนมากขึ้น และสามารถสื่อสารได้หลากหลายรูปแบบ เช่น วีดิโอ ภาพเสียง เกม หรือเนื้อหาที่น่าสนใจ นอกเหนือจากการสอนหน้าชั้นเรียนเพียงอย่างเดียว รวมถึงทำให้ผู้เรียนมีความรู้และทักษะด้านเทคโนโลยีสารสนเทศมากขึ้น และทำให้ผู้เรียนจดจำต่อสิ่งที่สนใจทำให้เกิดสมาธิในการเรียนรู้หรือจดจำในเรื่องใดเรื่องหนึ่ง โดยพยายามดึงความสนใจของลูก ให้มีความสุข สนุก สมาธิและจดจำอยู่กับของสิ่งนั้น

ตั้งแต่เริ่มแรกผ่านทุกๆ ประสาทสัมผัส และเค็ๆ จะเก็บความจำต่อเรื่องต่างๆ ได้อย่างมีประสิทธิภาพมากขึ้น

จากทฤษฎีการจัดการเรียนรู้เพื่อให้เกิดประโยชน์กับผู้เรียนนั้น องค์ประกอบที่สำคัญที่สุดคือ ครูผู้สอนที่จะจัดการเรียนการสอนเพื่อให้ประสบผลสำเร็จ นักเรียนได้รับความรู้ เกิดความเข้าใจและสามารถนำไปใช้ประโยชน์ในการดำรงชีวิตได้ ครูผู้สอนจึงจำเป็นต้องพัฒนารูปแบบการสอน กิจกรรมการเรียนการสอนเพื่อให้นักเรียนมีส่วนร่วมในกิจกรรมและมีความสุขในการเรียนมากที่สุด ซึ่งสอดคล้องกับกระทรวงศึกษาธิการ สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน เรื่อง การใช้หลักสูตรแกนกลางการศึกษาขั้นพื้นฐานโดยกำหนดให้สถานศึกษาในสังกัดจัดการเรียนการสอนโดยใช้หลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 [1] ในชั้นประถมศึกษาปีที่ 1-6 แต่สำหรับการเรียนการสอนในรายวิชาคอมพิวเตอร์ของกลุ่มสาระการเรียนรู้การงานอาชีพและเทคโนโลยี ชั้นประถมศึกษาปีที่ 2 ของโรงเรียนวัดลาดปลาเค้า จังหวัดนครปฐม ได้นำหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 มาใช้พัฒนานักเรียนให้มีคุณภาพตามมาตรฐานการเรียนรู้และมุ่งเน้นให้นักเรียน มีความรู้อันเป็นสากล เกิดสมรรถนะสำคัญได้แก่ความสามารถในการสื่อสาร การคิด การแก้ปัญหา การใช้ทักษะชีวิต และการใช้ เทคโนโลยี รวมทั้งเสริมสร้างนักเรียนให้มีคุณลักษณะ อันพึงประสงค์ สามารถอยู่ร่วมกับผู้อื่นในสังคมไทยและสังคมโลกได้อย่างมีความสุข ซึ่งการสอนรายวิชา คอมพิวเตอร์ของนักเรียนชั้นประถมศึกษาปีที่ 2 ของโรงเรียน โรงเรียนวัดลาดปลาเค้า จังหวัดนครปฐม นั้น มีรายละเอียดเนื้อหาวิชาดังนี้ [2] เรื่อง ส่วนประกอบของคอมพิวเตอร์ และในการจัดการเรียนปัญหาที่พบมีหลายประการเช่น 1) เนื้อหาในรายวิชามีคำศัพท์ที่จำยากและนักเรียนไม่เข้าใจจึงต้องนำรูปภาพมาแสดง 2) เวลาในการจัดการเรียนการสอน เนื่องจากข้อจำกัดของห้องคอมพิวเตอร์ ทำให้เครื่องคอมพิวเตอร์ไม่เพียงพอต่อนักเรียน 3) แอปพลิเคชันที่มาพร้อมกับอุปกรณ์แท็บเล็ตสำหรับนักเรียนชั้นประถมศึกษาปีที่ 2 นั้นที่โรงเรียนยังไม่สามารถใช้งานผ่านอินเทอร์เน็ตได้ [3] ซึ่งปรากฏดังนี้

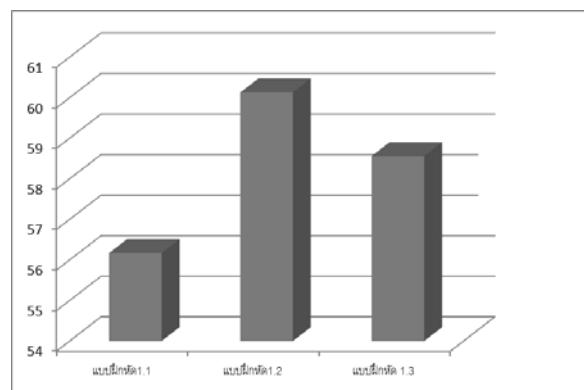


ภาพที่ 1: วิชาเรียนภายในแอปพลิเคชันในโครงการ OTPC

(One Tablet Per Child)

4) ผลสัมฤทธิ์ทางการเรียนของนักเรียนต่ำ ดังตารางที่ 1 ตารางที่ 1 แสดงจำนวนคะแนนแบบฝึกหัดท้ายบทเรียนเรื่องส่วนประกอบของคอมพิวเตอร์

แบบฝึกหัด1.1	แบบฝึกหัด1.2	แบบฝึกหัด1.3
56.17	60.13	58.56



ภาพที่ 2: จำนวนคะแนนแบบฝึกหัดท้ายบทเรียนเรื่องส่วนประกอบของคอมพิวเตอร์

ดังนั้นผู้วิจัยจึงมีแนวคิดในการพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนแบบเกม 3 มิติในรายวิชาคอมพิวเตอร์ ชั้นประถมศึกษาปีที่ 2 กรณีศึกษาโรงเรียนวัดลาดปลาเค้า โดยนำเนื้อหาประเภทเกม 3 มิติ ทำให้เกิดความน่าสนใจ และเอื้อต่อการเรียนรู้ของนักเรียน สามารถเรียนรู้ได้ตามความต้องการของ

แต่ละบุคคล เกิดกระบวนการเรียนรู้ได้เร็วขึ้น เกิดทักษะการเรียนรู้อย่างต่อเนื่องในการจดจำ

1.2 วัตถุประสงค์ของการวิจัย

1.2.1 เพื่อนำเสนอแนวทางการพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนแบบเกม 3 มิติในรายวิชาคอมพิวเตอร์ ชั้นประถมศึกษาปีที่ 2 กรณีศึกษาโรงเรียนวัดลาดปลาเค้า

1.2.2 เพื่อศึกษาความพึงพอใจของผู้ใช้บทเรียนคอมพิวเตอร์ช่วยสอนแบบเกม 3 มิติ

1.2.3 เพื่อศึกษาผลสัมฤทธิ์ทางการเรียนด้วยบทเรียนคอมพิวเตอร์ช่วยสอนแบบเกม 3 มิติ

1.3 สมมติฐานของการวิจัย

ประสิทธิภาพของบทเรียนคอมพิวเตอร์ช่วยสอนแบบเกม 3 มิติ ในรายวิชาคอมพิวเตอร์ ชั้นประถมศึกษาปีที่ 2 มีค่าสูงกว่าเกณฑ์ 80/80

2. วรรณกรรมที่เกี่ยวข้อง

ปัจจุบันมีการใช้บทเรียนคอมพิวเตอร์ช่วยสอน (CAI) เพื่อใช้งานผ่านเครือข่ายอินเทอร์เน็ต ซึ่งสามารถใช้งานด้วยเว็บเบราว์เซอร์ [4] ถือเป็นสื่อมัลติมีเดียเพื่อการศึกษาที่เปิดโอกาสให้ผู้เรียนเข้าถึงข่าวสารผ่านคอมพิวเตอร์ได้ง่าย สามารถตอบโต้และปฏิสัมพันธ์กับสื่อโดยตรง เมื่อมีการนำสื่อมัลติมีเดียมาใช้ในการศึกษา จึงเรียกว่า “สื่อมัลติมีเดียเพื่อการศึกษา” [5]

จากการศึกษาพบว่ามีผู้วิจัยที่ได้จัดทำสื่อ ในรูปของการทดลองเสมือนจริง ดังต่อไปนี้ [6] ได้ทำการวิจัยเรื่อง การพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนแบบจำลองสถานการณ์ บนเครือข่ายอินเทอร์เน็ตเรื่อง กฎหมายจราจร โดยใช้การจำลองสถานการณ์ที่สร้างขึ้นในคอมพิวเตอร์ เพื่อช่วยแก้ปัญหา การขาดสื่อการสอน โดยใช้การจำลองสถานการณ์ที่สร้างขึ้นในคอมพิวเตอร์มีความพึงพอใจต่อบทเรียนอยู่ในระดับมากที่สุด ซึ่งสอดคล้องกับ Barbara A. Sommer, Robert Sommer [7] ที่ได้พัฒนาบทเรียนคอมพิวเตอร์บน เครือข่ายแบบปรับเปลี่ยนเนื้อหาโดยนำเสนอในรูปแบบการ์ตูน 2 มิติ วิชา คณิตศาสตร์คอมพิวเตอร์ หลักสูตรประกาศนียบัตร วิชาชีพ สำนักงานคณะกรรมการการอาชีวศึกษา ร่วมกับวิชา การศึกษาเปรียบเทียบผลสัมฤทธิ์ทางการเรียนและความสนใจ ของนักเรียนช่วงชั้นที่ 2 ที่มีต่อการ์ตูนแอนิเมชัน 2 มิติและ 3 มิติ ของ นอกจากนี้ยังพบว่านักเรียนมีความสนใจต่อการ์ตูนแอนิเมชัน 2 มิติและ 3 มิติอยู่ในระดับ ดีและ

ผลสัมฤทธิ์ทางการเรียนของกลุ่มที่เรียนด้วยการ์ตูนแอนิเมชัน 2 มิติ และ 3 มิติ ไม่มีความแตกต่างกัน และ Bobbert ได้ศึกษาเกี่ยวกับประสิทธิภาพของการใช้ คอมพิวเตอร์จำลองแบบการทดลองวิชาเคมี กับการเรียนด้วยการทดลองด้วยตนเอง ตัวอย่างประชากรที่ลงทะเบียน วิชาเคมี จำนวน 153 คน โดยแบ่งประชากรออกเป็น 3 กลุ่ม คือ กลุ่มที่ใช้วิธีการเรียนการสอนแบบปกติ กลุ่มที่ใช้แบบจำลอง คอมพิวเตอร์ และกลุ่มที่ใช้ทั้งสองแบบ ผลการวิจัยพบว่าการ ทดสอบสสาร โดยการใช้แบบจำลองคอมพิวเตอร์มี ประสิทธิภาพด้านการเรียนการสอนเหมือนกับการสอนแบบ ปกติ สำหรับการทดลองเรื่องกฎแห่งก๊าซของบอยล์ นักศึกษาที่ทำการทดลองด้วยตนเอง และมีประสบการณ์กับแบบจำลองคอมพิวเตอร์ จะได้คะแนนมากกว่าอย่างมีนัยสำคัญทางสถิติ ระหว่างกลุ่มนักเรียนที่ทำการทดลองด้วยตนเองอย่างเดียว นอกจากนั้นนักเรียนยังให้ความเห็นว่าเขาสนใจที่จะ ปฏิสัมพันธ์กับการแบบจำลองคอมพิวเตอร์ที่มีวิธีการเรียน เป็นที่พึงพอใจ ดังนั้นผู้วิจัยจึงได้คิดพัฒนาบทเรียน เพื่อสื่อ ประโยชน์ต่อผู้เรียนเพื่อที่จะค้นคว้าหาความรู้ได้ตามความต้องการซ้ำเร็วของแต่ละบุคคลที่ไม่เหมือนกัน ได้และการที่ ผู้เรียนสามารถควบคุมความซ้ำเร็วได้ตามความพอใจ นั้นจะไม่ทำให้เกิดความเครียดในการเรียน และมีทัศนคติที่ดีขึ้นต่อครูผู้สอนและครูผู้สอนยังสามารถเก็บผลการศึกษานักเรียนได้ ทำให้เกิดทักษะการเรียนรู้อย่างต่อเนื่อง และเกิดความรวดเร็วในการเรียนการสอนซึ่งเป็นแรงจูงใจดึงดูดให้นักเรียนอยากเรียนมากขึ้น และเรียนรู้ครบตามหลักสูตรมากยิ่งขึ้น

3. วิธีการดำเนินการวิจัย

3.1 กรอบแนวคิดในการวิจัย

3.1.1 กรอบแนวคิดด้านสื่อ

ในการออกแบบการพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนแบบเกม 3 มิติในรายวิชาคอมพิวเตอร์ ชั้นประถมศึกษาปีที่ 2 กรณีศึกษาโรงเรียนวัดลาดปลาเค้าที่มีประสิทธิภาพ โดยผู้วิจัยได้ใช้วิธี [8] เป็นกรอบแนวคิดในการสร้างบทเรียนคอมพิวเตอร์ช่วยสอนโดยมีขั้นตอน ดังนี้ 1) เตรียมข้อมูลเครื่องมือ 2) โครงร่างหน้าจอ 3) สอบถามผู้เชี่ยวชาญ 4) เขียนบทดำเนินเรื่อง 5) ออกแบบผังงาน 6) สร้างบทเรียนต้นแบบ

3.1.2 กรอบแนวคิดด้านเนื้อหา

ในการออกแบบเนื้อหาครั้งนี้ ส่วนประกอบของคอมพิวเตอร์

3.1.3 กรอบแนวคิดด้านการวัดและประเมิน

ผู้วิจัยนำแนวคิดการสร้างบทเรียนคอมพิวเตอร์ช่วยสอนซึ่งใช้เทคนิคการออกแบบการเรียนการสอนที่ดัดแปลงมาจากการออกแบบระบบการสอน โดยใช้แบบจำลอง ADDIE Model [9] ประกอบด้วยขั้นตอนต่อไปนี้ 1) การวิเคราะห์ 2) การออกแบบ 3) การพัฒนา 4) การนำไปใช้ 5) การประเมินผล

3.2 ขอบเขตของการวิจัย

กลุ่มตัวอย่างที่ใช้การวิจัยครั้งนี้คือนักเรียนชั้นประถมศึกษาปีที่ 2 โรงเรียนวัดลาดปลาเค้า ภาคเรียนที่ 1/2556 ได้มาด้วยวิธีการเลือกแบบสุ่มอย่างง่ายจำนวน 30 คน

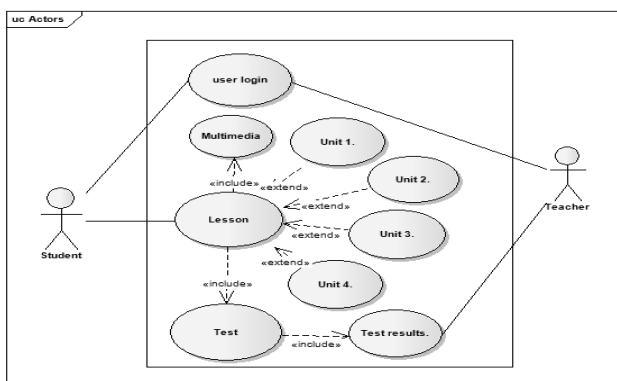
3.3 เครื่องมือที่ใช้ในการวิจัย

3.3.1 บทเรียนคอมพิวเตอร์ช่วยสอนแบบเกม 3 มิติในรายวิชาคอมพิวเตอร์

3.3.1.1 ดำเนินการสร้างบทเรียนคอมพิวเตอร์ช่วยสอนแบบเกม 3 มิติในรายวิชาคอมพิวเตอร์ มีขั้นตอนดังต่อไปนี้

- ก) รวบรวมเนื้อหาข้อมูลที่ได้จากการศึกษาเนื้อหาบทเรียนของนักเรียนชั้นประถมศึกษาปีที่ 2
- ข) สร้างบทเรียนโดยใช้โปรแกรมอื่นๆ ที่เกี่ยวข้อง ดังนี้

- โปรแกรม 3D Game Maker
- โปรแกรม Photoshop
- ค) พัฒนบทเรียนตามโครงสร้างต่อไปนี้



ภาพที่ 3: ภาพรวมการทำงานของแอปพลิเคชัน

3.3.2 แบบประเมินความพึงพอใจของบทเรียนคอมพิวเตอร์ช่วยสอนแบบเกม 3 มิติ

3.3.3 แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน ที่มีลักษณะแบบทดสอบปรนัย 3 ตัวเลือกจำนวน 10 ข้อ

4. ผลการดำเนินงาน

เมื่อนำบทเรียนที่พัฒนาไปใช้ สามารถเข้าถึงเนื้อหาหน่วยการเรียนรู้ได้ครบถ้วน และสามารถนำเสนอในรูปแบบของสื่อมัลติมีเดียได้ทั้งในส่วนของภาพ เสียง วิดีโอ และการประเมินด้วยแบบทดสอบ ดังภาพที่ 4-7



ภาพที่ 4 : แสดงการเข้าสู่บทเรียนคอมพิวเตอร์ช่วยสอนแบบเกม 3 มิติ



ภาพที่ 5 : แสดงตัวอย่างการออกแบบหน้าจอตรวจสอบคะแนน

4.1 ผลการประเมินความพึงพอใจของบทเรียนคอมพิวเตอร์ช่วยสอนแบบเกม 3 มิติ แสดงดังตารางที่ 2

ตารางที่ 2 ผลการประเมินความพึงพอใจของบทเรียนคอมพิวเตอร์ช่วยสอนแบบเกม 3 มิติ

รายการประเมิน	ค่าเฉลี่ย	S.D.	ระดับคุณภาพ
1. ด้านเนื้อหา	4.15	0.62	ดี
2. ด้านตัวอักษรและสี	4.56	0.66	ดีมาก
3. การจัดการบทเรียน	4.23	0.62	ดี
4. การปฏิสัมพันธ์กับบทเรียน	4.65	0.58	ดีมาก
รวม	4.39	0.62	ดี

จากตารางที่ 2 แสดงให้เห็นว่าผลจากการประเมินความพึงพอใจของนักเรียนอยู่ในระดับดี โดยมีค่าเฉลี่ย 4.39 และค่าเบี่ยงเบนมาตรฐาน 0.62 ตามลำดับ

4.2 ผลการหาประสิทธิภาพของบทเรียนคอมพิวเตอร์ช่วยสอนแบบเกม 3 มิติ

จากการทำแบบทดสอบเพื่อหาประสิทธิภาพของบทเรียนแสดงดังตารางที่ 3

ตารางที่ 3 ผลการหาประสิทธิภาพของบทเรียนคอมพิวเตอร์ช่วยสอนแบบเกม 3 มิติ

ทดสอบ	ร้อยละ
คะแนนระหว่างเรียน (E1)	85.21
คะแนนสอบหลังเรียน (E2)	80.06

จากตารางที่ 3 แสดงให้เห็นว่าบทเรียนคอมพิวเตอร์ช่วยสอนแบบเกม 3 มิติ มีประสิทธิภาพ (E1/E2) เท่ากับ 85.21/80.06 ซึ่งเป็นไปตามเกณฑ์ที่กำหนด

5. สรุปและอภิปรายผล

5.1 สรุปผลการวิจัย

การพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนแบบเกม 3 มิติ ในรายวิชาคอมพิวเตอร์ ของชั้นประถมศึกษาปีที่ 2 นั้น สามารถนำไปใช้ในการเรียนการสอนได้จริง และทำให้นักเรียนเกิดความสนใจในวิชาที่เรียนมากขึ้น พิจารณาได้จากผลการประเมินความพึงพอใจของครูผู้สอนและนักเรียนที่มีต่อบทเรียน นอกจากนี้เป็นการประยุกต์ใช้เทคโนโลยีให้เกิดประโยชน์สูงสุดย่อมส่งผลต่อการพัฒนาอย่างยั่งยืนสืบไป

5.2 ข้อเสนอแนะในการวิจัยครั้งต่อไป

- 5.3.1 ควรมีการพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนแบบเกม 3 มิติ ในรายวิชาอื่นๆได้
- 5.3.2 ควรมีการเพิ่ม interactive ในบทเรียนให้มากขึ้น
- 5.3.3 ควรมีการสร้างการ์ตูนเคลื่อนไหวให้มากยิ่งขึ้น

6. เอกสารอ้างอิง

- [1] กระทรวงศึกษาธิการ. *หลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551*. กรุงเทพฯ: สำนักวิชาการและมาตรฐานการศึกษา สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน, 2551.
- [2] อำนวย เดชชัยศรีและคณะ, "เทคโนโลยีสารสนเทศ ชั้นประถมศึกษาปีที่ 1" สำนักพิมพ์วัฒนาพานิช จำกัด, 2553
- [3] OTPC (One Tablet Per Child), "โครงการ OTPC" สืบค้นวันที่ 2 ธันวาคม 2555, [ออนไลน์] <http://www.dekthai.net/lo.html>
- [4] ถนอมพร (ต้นพิพัฒน์) เลหาจรัสแรง. "คอมพิวเตอร์ช่วยสอน" กรุงเทพฯ.ภาควิชาโสตทัศนศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2541.
- [5] สื่อมัลติมีเดียเพื่อการศึกษา, "ความหมายของสื่อมัลติมีเดียเพื่อการศึกษา" สืบค้นวันที่ 2 ธันวาคม 2553, [ออนไลน์] http://cmi.nfego.th/multimedialesson/multi_lesson/lesson/01/meaning.html
- [6] สวียา สุรมณี. "การพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนแบบจำลองสถานการณ์บนเครือข่ายอินเทอร์เน็ต เรื่อง กฎหมายจราจร" มหาวิทยาลัยราชภัฏมหาสารคาม, 2552.
- [7] Barbara A. Sommer, Robert Sommer. "A Virtual Lab in ResearchMethods" University of California ,Daris , 2003.
- [8] เฉลิมชัย ดาระกา, "การพัฒนาบทเรียนคอมพิวเตอร์บนเครือข่ายแบบปรับเปลี่ยนเนื้อหาโดยนำเสนอในรูปแบบการ์ตูน 2 มิติวิชาคณิตศาสตร์คอมพิวเตอร์ หลักสูตรประกาศนียบัตรวิชาชีพ สำนักงานคณะกรรมการการอาชีวศึกษา. สาขาเทคโนโลยีคอมพิวเตอร์". มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ,2551
- [9] ธงชัย วิไลวิทย์. "ADDIE Model". สืบค้นเมื่อวันที่ 17 มกราคม 2553, จาก http://student.nu.ac.th/comed/webboard/answer_questID=6

การพัฒนาสื่อประกอบการสอนแบบมีส่วนร่วมออนไลน์ด้วยการเข้าถึงข้อมูล
ผ่านบาร์โค้ด 2 มิติ เพื่อวัดความสนใจในเนื้อหาของผู้เรียนและ
เสริมทักษะการฝึกปฏิบัติใช้งานโปรแกรมสำเร็จรูป

**Development of Online Participatory Teaching Media Access via 2D Barcode For Measure
of Intention in Learning Context and Enhancing Practice Package Software Skill**

อุบลรัตน์ ศิริสุข โภคา (Ubonrat Sirisukpoca)¹ และ ไพศาล สิมิลาเต่า (Paisan Simalaotao)²

^{1,2} กลุ่มสาขาวิชาคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครปฐม

บทคัดย่อ

บทความนี้เป็นส่วนหนึ่งของงานวิจัยเพื่อแก้ปัญหาทางการศึกษาขาดทักษะในการฝึกปฏิบัติการใช้งาน โปรแกรมสำเร็จรูป รายวิชา เทคโนโลยีสารสนเทศและการสื่อสารด้วยการใช้สื่อประกอบการสอนแบบมีส่วนร่วมเพื่อช่วยในการฝึกปฏิบัติ มีวัตถุประสงค์เพื่อ 1) พัฒนาสื่อประกอบการสอนแบบมีส่วนร่วมออนไลน์ที่เข้าถึงข้อมูลผ่านบาร์โค้ด 2 มิติ 2) เพื่อเสริมทักษะในการฝึกปฏิบัติการใช้งาน โปรแกรม โดยกำหนดประชากรที่ใช้ในการศึกษาเป็นนักศึกษาปริญญาตรีจำนวน 70 คน เครื่องมือที่ใช้ในการวิจัย คือ สื่อประกอบการสอนแบบมีส่วนร่วมออนไลน์ที่สามารถเข้าถึงได้ผ่านบาร์โค้ด 2 มิติ ขั้นตอนการวิจัยเป็นการออกแบบและพัฒนาสื่อประกอบการสอนและทดลองให้ผู้ใช้งานเข้าถึงด้วยเครื่องคอมพิวเตอร์หรืออุปกรณ์พกพาชนิดต่าง ๆ ผ่านบาร์โค้ด 2 มิติ แล้วทำการเก็บข้อมูลการใช้งานสื่อประกอบการสอนของผู้ใช้ตลอดการใช้งาน เพื่อประเมินความสนใจในเนื้อหาของผู้ใช้ ผลการดำเนินงานใช้การประเมินคุณภาพของสื่อประกอบการสอนและระบบจากผู้เชี่ยวชาญจำนวน 2 ท่าน อยู่ในระดับดีมากโดยมีค่าเฉลี่ยอยู่ที่ 4.60 และค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.50 และผลการประเมินความพึงพอใจจากผู้ใช้งานจำนวน 70 คน อยู่ในระดับดีมาก โดยมีค่าเฉลี่ยอยู่ที่ 4.72 และค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.45 ตามลำดับ สรุปได้ว่าสื่อประกอบการสอนแบบมีส่วนร่วมออนไลน์ด้วยการเข้าถึงข้อมูลผ่านบาร์โค้ด 2 มิติ สามารถใช้งานได้ถูกต้องตามหลักวิชาการและเป็นที่ยอมรับของผู้ใช้เป็นอย่างมาก และช่วยเพิ่มทักษะการใช้งาน โปรแกรมสำเร็จรูปได้

คำสำคัญ: สื่อประกอบการสอนแบบมีส่วนร่วม บาร์โค้ด 2 มิติ

Abstract

This article is part of the research to solve students lack the skills to practice use application of Information and Communication Technology courses by using involved teaching materials for practical activities. The purpose were to 1) develop participatory teaching media online access via 2D barcode 2) to enhance skills in the practice application. By the population used in the study were 70 undergraduates. The tools used in this research is teaching participatory media online can be reached via 2D barcodes. Research phase to design and develop the teaching and testing a user to access a computer or portable device types via 2D barcodes and then perform data access Teaching media of users throughout the work To assess the interest in the contents of the user. Performance evaluation of the quality of the Teaching media professionals to 2 people, is very good, with an average of 4.60 and a standard deviation of 0.50 and the satisfaction of 70 users is very good. with an average is 4.72 and the standard deviation was 0.45, respectively. Concluded that media Teaching participatory online by accessing data through a 2D barcode can be used as a legal academic, and user satisfaction in a very good level and enhances the skills to use application software.

Keyword: Teaching participatory media, Teaching media, 2D barcodes.

1. บทนำ

ส่วนหนึ่งของการเรียนการสอนในระดับมหาวิทยาลัยที่มีความจำเป็นต่อการสำเร็จการศึกษาประกอบไปด้วยความรู้ทางด้านภาษาอังกฤษและความสามารถทางด้านเทคโนโลยีสารสนเทศ ซึ่งเป็นการประยุกต์ใช้งานซอฟต์แวร์สำเร็จรูปในสำนักงาน จึงได้จัดให้มีการเรียนการสอนรายวิชา เทคโนโลยีสารสนเทศและการสื่อสาร ให้กับนิสิตและนักศึกษาทุกสาขาวิชา ทั้งการบรรยายเนื้อหาทางทฤษฎีและการฝึกปฏิบัติ แต่เนื่องจากนักศึกษาแต่ละสาขาวิชามีความสามารถพื้นฐานในการใช้งานคอมพิวเตอร์แตกต่างกัน และมีช่วงเวลาในการทบทวนในเวลาเรียนอย่างจำกัด

ผู้วิจัยจึงมีแนวคิดในการพัฒนาสื่อประกอบการสอนแบบมีส่วนร่วมออนไลน์ด้วยการเข้าถึงข้อมูลผ่านบาร์โค้ด 2 มิติเพื่อวัดความสนใจในเนื้อหาของผู้เรียนและเสริมทักษะการฝึกปฏิบัติใช้งานโปรแกรมสำเร็จรูปทั้งในเวลาเรียนและนอกเวลาเรียน

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องประกอบด้วยงานวิจัยทางด้านสื่อการสอนและงานวิจัยทางด้านเทคโนโลยีบาร์โค้ด 2 มิติ (QR Code) ซึ่งมีงานวิจัยที่เกี่ยวข้องมากมาย เช่น การพัฒนารูปแบบแบบเรียนฝึกความสามารถผ่านเครือข่ายอินเทอร์เน็ต เรื่องการตัดต่อเสียงสำหรับการสร้างสื่อประกอบการสอนวิทยาศาสตร์ ซึ่งเป็นงานวิจัยที่มีการใช้สื่อการสอนบนผ่านเครือข่ายอินเทอร์เน็ต [1] งานวิจัยการศึกษาผลสัมฤทธิ์ของการสอนแบบในชั้นเรียนกับการสอน โดยใช้สื่อการสอนผ่านเครือข่ายคอมพิวเตอร์ กรณีศึกษา รายวิชา การเขียน โปรแกรมคอมพิวเตอร์ 1 [2] พบว่า ผู้เรียนมีความเข้าใจเนื้อหาเพิ่มมากขึ้น และได้ผลสัมฤทธิ์ทางการเรียนสูงขึ้นกว่าการสอนในชั้นเรียนอย่างเดียวน การผลิตสื่อการสอนประกอบรายวิชาคอมพิวเตอร์และเทคโนโลยีสารสนเทศพื้นฐาน [3] พบว่าผู้เรียนมีความพึงพอใจในการใช้สื่อการสอนเพิ่มสูงขึ้น งานวิจัยการพัฒนาสื่อการสอนบนเว็บ เรื่องการสื่อสารข้อมูลและเครือข่ายคอมพิวเตอร์ [4] พบว่า สื่อการสอนบนเว็บช่วยให้ผู้เรียนเกิดความสนใจและสามารถเข้าถึงแหล่งเรียนรู้ได้ตลอดเวลาจนทำ

ให้ผู้เรียนเกิดความพึงพอใจในระดับดี การพัฒนาซอฟต์แวร์ระบบฐานข้อมูลบุคลากรและนักเรียน ในโรงเรียนประถมศึกษาโดยใช้บาร์โค้ด [5] พบว่า การใช้บาร์โค้ดช่วยลดขั้นตอนการทำงานได้และผู้ใช้ทุกระดับเกิดความพึงพอใจในระดับดี การประยุกต์ใช้คิวอาร์โค้ดระบุตำแหน่งด้วยเครือข่ายโทรศัพท์มือถือ และระบบฐานข้อมูลสำหรับผู้โดยสารอากาศยานขาเข้าที่โดยสารรถแท็กซี่สาธารณะสังกัดท่าอากาศยานสุวรรณภูมิ [6] พบว่าการใช้เทคโนโลยีคิวอาร์โค้ดช่วยลดขั้นตอนการนำเข้าข้อมูลที่จำเป็นได้และผู้ใช้ได้รับความพึงพอใจในระดับดี นอกจากนี้ยังมีงานวิจัยอื่น ๆ ที่เกี่ยวข้อง เช่น ระบบสารสนเทศเพื่อการบริหารครุภัณฑ์คอมพิวเตอร์ด้วย Quick Response Code กรณีศึกษา สำนักงานการตรวจเงิน [7] การประยุกต์ใช้บาร์โค้ดสองมิติชนิดคิวอาร์เพื่อการจัดการ [8] และการประยุกต์ใช้บาร์โค้ดสองมิติชนิดคิวอาร์เพื่อการจัดการร้านค้า [9] เป็นต้น

จากงานวิจัยที่กล่าวมาข้างต้น สรุปได้ว่า การพัฒนาสื่อประกอบการสอนที่ให้บริการผ่านเครือข่ายอินเทอร์เน็ตทำให้ผู้เรียนเกิดความพึงพอใจในการใช้งานและช่วยให้ผลสัมฤทธิ์ทางการเรียนสูงขึ้น และการเข้าถึงข้อมูลผ่านเทคโนโลยีบาร์โค้ด 2 มิติ ทำให้เกิดความสะดวกรวดเร็วในการเข้าถึงข้อมูล ผู้วิจัยจึงได้นำเทคโนโลยีบาร์โค้ด 2 มิติ มาช่วยเพิ่มประสิทธิภาพในการเข้าถึงสื่อประกอบการสอนออนไลน์

2.2 ทฤษฎีที่เกี่ยวข้อง

เทคโนโลยีบาร์โค้ดแบบ 2 มิติ (QR Code)

คุณสมบัติของ QR CODE

1. สามารถบรรจุข้อมูลได้ในปริมาณสูง ทั้งตัวอักษรเลข (Alphanumeric) ตัวอักษรภาษาญี่ปุ่น (ทั้ง Kanji และ Hiragana) สัญลักษณ์ ตัวเลขฐานสอง (binary) และรหัสสี (colure code) โดยทั้งหมดนี้สามารถจะบรรจุไว้ได้ในคราวเดียวกัน



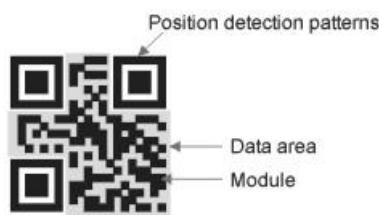
ภาพที่ 1: การบรรจุข้อมูลของ Alphanumeric จำนวน 300 ตัว ไว้ใน QR CODE เพียง 1 ภาพ

2. สามารถบันทึกข้อมูลได้ทั้งในแนวตั้งและแนวนอน ทำให้ใช้พื้นที่ในการบันทึกน้อยกว่าบาร์โค้ดปกติ

3. ความสามารถในการบรรจุตัวอักษรภาษาญี่ปุ่น ทำให้ QR CODE ได้รับ Japanese Industrial Standard (JIS) หรือมาตรฐานอุตสาหกรรมของประเทศญี่ปุ่นซึ่งเป็นสัญลักษณ์บอกว่า QR CODE นี้สามารถใช้ได้ในกิจกรรมอุตสาหกรรมในประเทศญี่ปุ่น ซึ่งการบรรจุข้อมูลสูงสุดถึง 13 bits (26 ตัวอักษร) ซึ่งมากกว่า 2D CODE แบบอื่นถึง 20 %

4. ป้องกันการสกรปรกและการฉีกขาด สามารถอ่านข้อมูลหรือกู้ข้อมูลได้แม้ว่าจะมีการฉีกขาดหรือมีการ สกรปรก เพียงบางส่วน โดยสามารถกู้คืนได้มากที่สุด 30% ของ CODEWORD (1 codeword= 8 bits หรือ 16 ตัวอักษร)

5. สามารถอ่านข้อมูลได้ 360 องศา ด้วยความเร็วสูง ผ่านรูปแบบของการตรวจสอบตำแหน่ง ที่อยู่ทั้ง 3 มุมของ สัญลักษณ์ซึ่งรูปแบบการตรวจสอบเหล่านี้ทำให้เครื่องอ่านมีความเสถียร ในเรื่องของความเร็วในการอ่าน และเป็นตัวป้องกันการรบกวนของพื้นหลัง



ภาพที่ 2: รูปแบบของการตรวจสอบตำแหน่ง ที่อยู่ทั้ง 3 มุม

3. วิธีดำเนินการวิจัย

1. ศึกษาเบื้องต้น (Preliminary Investigation and Feasibility Study) เกี่ยวกับการสร้างสื่อประกอบการสอนออนไลน์ การสร้างและอ่านบาร์โค้ด 2 มิติ และการพัฒนาเว็บแอปพลิเคชัน

ผลการศึกษาพบว่ามีความเป็นไปได้ในการประยุกต์ใช้สร้างเครื่องมือสำหรับใช้ในงานวิจัยได้เป็นอย่างดี

2. ศึกษากระบวนการทำงานและกำหนดปัญหา (Problem Specification) จากการศึกษาถึงปัจจัย ดังต่อไปนี้

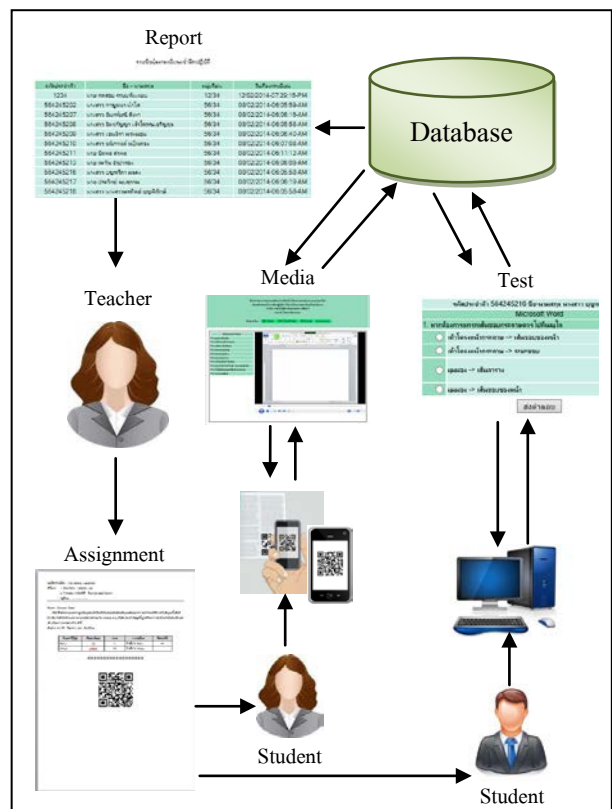
2.1 ศึกษาถึงปัจจัยที่ทำให้นักศึกษาขาดทักษะในการฝึกปฏิบัติการใช้งานซอฟต์แวร์สำเร็จรูปที่ใช้ในสำนักงาน

2.2 ศึกษาและรวบรวมความต้องการของผู้เรียนเกี่ยวกับช่องทางการเข้าถึงสื่อประกอบการสอน

2.3 ศึกษาถึงการพัฒนาสื่อประกอบการสอนให้สอดคล้องกับเนื้อหาในบทเรียน

ผลการศึกษาพบว่าปัจจัยที่ทำให้ให้นักศึกษาขาดทักษะในการฝึกปฏิบัติ คือ ขาดสื่อประกอบการสอนสำหรับใช้เป็นแหล่งหาความรู้และมีความต้องการฝึกทักษะนอกเวลาเรียน โดยช่องทางที่ช่วยให้สะดวกในการเข้าถึงสื่อ ได้แก่ เว็บไซต์บนเครือข่ายอินเทอร์เน็ต และการใช้งานผ่านสมาร์ทโฟน ซึ่งสื่อประกอบการสอนที่ทำให้เกิดความสนใจ ได้แก่ สื่อประเภทวิดีโอ ภาพนิ่ง การ์ตูน และเกม เป็นต้น

จากผลการศึกษาทางผู้วิจัยจึงได้ออกแบบระบบตามสมมติฐาน เพื่อให้ผู้เรียนสามารถเข้าถึงสื่อประกอบการสอนได้หลากหลายช่องทาง ดังภาพที่ 3



ภาพที่ 3: การทำงานของระบบ

3. วิเคราะห์ระบบ (System Analysis)

3.1 วิเคราะห์แนวทางการพัฒนาสื่อประกอบการสอนแบบมีส่วนร่วม โดยเน้นรูปแบบการนำเสนอเนื้อหาจากผลการศึกษาเบื้องต้น โดยการแบ่งเนื้อหาออกเป็นหัวข้อย่อย และ

วิเคราะห์ความเป็นไปได้ในการสร้างสื่อประกอบการสอนที่มีความสอดคล้องกับเนื้อหาดังกล่าว

3.2 วิเคราะห์ช่องทางในการเข้าถึงข้อมูลผู้ใช้ ข้อมูลเนื้อหา ข้อมูลสื่อประกอบการสอน ข้อมูลผลการประเมินความสนใจของผู้เรียน และข้อมูลผลสัมฤทธิ์ทางการเรียน

4. ออกแบบระบบ (System Design)

4.1 ส่วนของการนำเข้าข้อมูล ประกอบด้วย ข้อมูลผู้เรียน ข้อมูลสื่อประกอบการสอน ข้อมูลงานที่ได้จากการฝึกปฏิบัติ ข้อมูลแบบทดสอบเพื่อวัดผลสัมฤทธิ์ทางการเรียน

4.2 ส่วนของการประเมินผล ประกอบด้วย การประเมินผลแบบทดสอบเพื่อวัดผลสัมฤทธิ์ทางการเรียน การประเมินความสนใจในเนื้อหาที่ผู้เรียนขาดความเข้าใจหรือให้ความสนใจ

4.3 ส่วนของการนำเสนอ ประกอบด้วย การใช้สื่อประกอบการสอนผ่านเครือข่ายอินเทอร์เน็ต การเข้าถึงข้อมูลเกี่ยวกับผู้เรียนหรือรายละเอียดเกี่ยวกับการฝึกปฏิบัติผ่านเอกสารติดบาร์โค้ด 2 มิติ ทำให้สามารถเข้าใช้งานผ่านอุปกรณ์พกพา การทำแบบทดสอบเพื่อวัดผลสัมฤทธิ์ทางการเรียนผ่านเว็บแอปพลิเคชัน

5. พัฒนาโปรแกรม (Coding)

5.1 พัฒนาเว็บแอปพลิเคชัน ประกอบด้วย ส่วนของระบบจัดการฐานข้อมูล การเข้าสู่ระบบ การประเมินผลแบบทดสอบเพื่อวัดผลสัมฤทธิ์ทางการเรียน การประเมินความสนใจในเนื้อหาที่ผู้เรียนขาดความเข้าใจหรือให้ความสนใจ ด้วยภาษา PHP เชื่อมต่อกับฐานข้อมูล MySQL

5.2 พัฒนาสื่อประกอบการสอน เช่น ภาพนิ่งและวิดีโอ ประกอบการสอนด้วย Adobe Photoshop และ Camtasia Studio

6. ทดสอบระบบ (Testing)

6.1 ทดสอบระบบส่วนย่อย ประกอบด้วย

1. ทดสอบการนำเข้าข้อมูลพื้นฐาน ข้อมูลสื่อประกอบการสอน และการส่งไฟล์งานภาคปฏิบัติผ่านเครือข่ายอินเทอร์เน็ต

2. ทดสอบความถูกต้องของผลการวิเคราะห์ข้อมูลเพื่อวัดผลสัมฤทธิ์ทางการเรียน และผลการประเมินความสนใจในเนื้อหาที่ผู้เรียนขาดความเข้าใจหรือให้ความสนใจในการเรียนรู้

3. ทดสอบการสร้างและอ่านบาร์โค้ด 2 มิติ

ผลการทดสอบระบบในแต่ละส่วนมีการทำงานที่ถูกต้อง สะดวก รวดเร็วและไม่พบปัญหาในการทำงาน การเข้าสู่ระบบผ่าน QR Code เพิ่มความสะดวกรวดเร็วในการใช้งานสื่อประกอบการสอนได้เป็นอย่างดี ผู้ใช้ไม่ต้องจดจำ URL ของเว็บเพจที่ให้บริการ

6.2 ทดสอบการทำงานร่วมกันของทั้งระบบ

7. นำเข้าข้อมูลจริง (Data Conversion & Data Entry) โดยการนำเข้าข้อมูลพื้นฐานและสื่อประกอบการสอนเข้าสู่ระบบพร้อมให้ผู้เรียนเข้าใช้งาน

8. เริ่มใช้งานกับข้อมูลจริง (Implementation) เปิดให้ผู้เรียนลงทะเบียนเข้าใช้งานจริง และทำการเก็บข้อมูลการใช้งาน

9. ปรับปรุงแก้ไข (Maintenance)

10. เผยแพร่ผ่านเครือข่ายอินเทอร์เน็ต

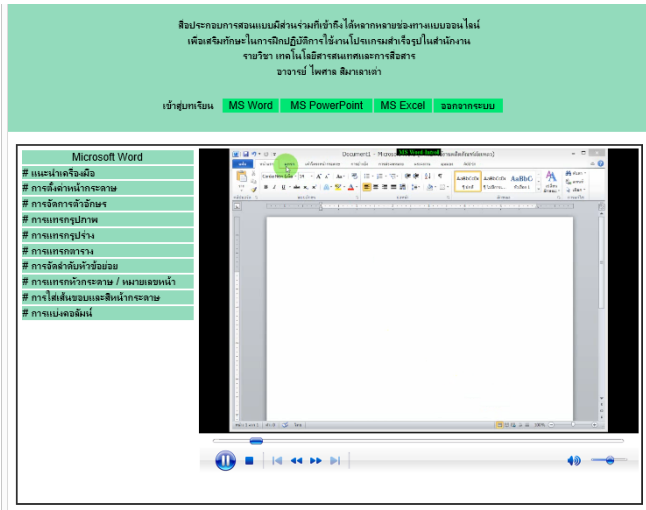
หลังทำการพัฒนาสื่อการสอนและนำเข้าข้อมูลจริงทั้งส่วนของข้อมูลผู้เรียนและข้อมูลเนื้อหารายวิชาแล้ว ทางผู้วิจัยได้เปิดให้ผู้เรียนเข้าใช้งานสื่อการสอน โดยมีรายละเอียด ดังนี้

10.1 ผู้เรียนสามารถเข้าถึงสื่อประกอบการสอนผ่านบาร์โค้ด 2 มิติ โดยใช้อุปกรณ์พกพาหรือเครื่องคอมพิวเตอร์ที่เชื่อมต่อกล้องถ่ายภาพแสดงดังภาพที่ 4



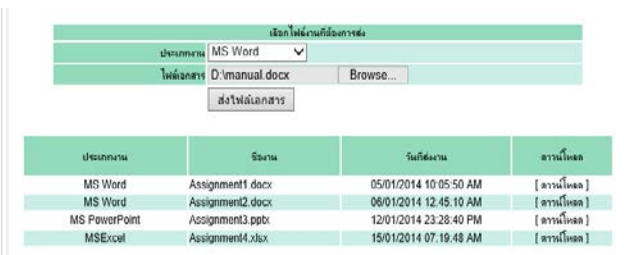
ภาพที่ 4: การเข้าถึงสื่อประกอบการสอนผ่านบาร์โค้ด 2 มิติ

10.2 ผู้เรียนสามารถใช้งานสื่อประกอบการสอน เพื่อเสริมทักษะการฝึกปฏิบัติการใช้งาน โปรแกรมสำเร็จรูป ด้วยสื่อชนิดต่าง ๆ เช่น เอกสารคู่มือการใช้งาน หรือไฟล์วิดีโอสาธิตการใช้งาน เป็นต้น ดังแสดงในภาพที่ 5



ภาพที่ 5: สื่อประกอบการสอนการฝึกปฏิบัติใช้งานโปรแกรมสำเร็จรูป

- ผู้เรียนสามารถส่งงานภาคปฏิบัติผ่านทางระบบ ดังแสดงในภาพที่ 6 เพื่อให้ผู้สอนตรวจพร้อมทั้งเรียกดูข้อมูลผลการประเมินความสนใจในเนื้อหา ดังแสดงในภาพที่ 7

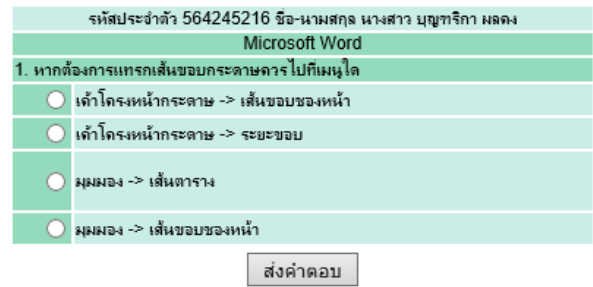


ภาพที่ 6: แสดงการส่งไฟล์งานภาคปฏิบัติ

รหัสประจำตัว 564245216 ชื่อ-นามสกุล นางสาว มยุษฎริกา ผลคง		
Microsoft Word	จำนวนครั้งที่เข้าใช้	
# แนะนำเครื่องมือ	1	
# การตั้งค่าหน้ากระดาษ	2	
# การจัดการตัวอักษร	1	
# การแทรกรูปภาพ	3	
# การแทรกตาราง	3	
# การแทรกตาราง	5	
# การจัดลำดับหัวข้อย่อย	3	

ภาพที่ 7: แสดงผลการประเมินความสนใจในเนื้อหาที่ผู้เรียนสนใจ

- ผู้เรียนสามารถเข้าทำแบบทดสอบวัดความเข้าใจเกี่ยวกับเนื้อหาของโปรแกรมสำเร็จรูปต่าง ๆ ดังแสดงในภาพที่ 8 และเรียกดูผลการวิเคราะห์ข้อมูลเพื่อวัดผลสัมฤทธิ์ทางการเรียนได้ดังแสดงในภาพที่ 9



ภาพที่ 8: แบบทดสอบเพื่อวัดผลสัมฤทธิ์ทางการเรียน

รหัสประจำตัว 564245216 ชื่อ-นามสกุล นางสาว มยุษฎริกา ผลคง			
Microsoft Office	Microsoft Word	Microsoft PowerPoint	Microsoft Excel
คะแนนสอบ	20	25	18
คะแนนงานภาคปฏิบัติ	8	10	7
จำนวนครั้งที่เข้าใช้ระบบ			
# แนะนำเครื่องมือ	1	1	1
# การตั้งค่าหน้ากระดาษ	1	1	1
# การจัดการตัวอักษร	1	0	0
# การแทรกรูปภาพ	1	0	0
# การแทรกตาราง	1	0	0
# การแทรกตาราง	1	1	1
# การจัดลำดับหัวข้อย่อย	1	0	0
# การแทรกหัวกระดาษ / หมายเลขหน้า	1	1	1
# การใส่เส้นขอบและสีหน้ากระดาษ	2	2	2
# การแนบคอมเมนต์	1	1	1

ภาพที่ 9: การวิเคราะห์ข้อมูลเพื่อวัดผลสัมฤทธิ์ทางการเรียน

4. ผลการดำเนินงาน

ผลการดำเนินงานใช้การประเมินคุณภาพของระบบโดยกลุ่มผู้เชี่ยวชาญด้านการพัฒนาระบบสารสนเทศจำนวน 2 ท่าน โดยมีเกณฑ์ให้คะแนนแบบ 5 คะแนน ดังแสดงในตารางที่ 1

ตารางที่ 1: ผลการประเมินคุณภาพของระบบจากผู้เชี่ยวชาญ

รายการประเมิน	ระดับคุณภาพ		
	X	S.D.	เชิงคุณภาพ
1. ด้านความสะดวกและง่ายต่อการใช้งานระบบ	4.75	0.43	ดีมาก
2. ด้านความสามารถตามหน้าที่ในการทำงานของระบบ	4.66	0.50	ดีมาก
3. ด้านความปลอดภัยของระบบ	4.33	0.57	ดี
4. ด้านความสามารถของระบบตามความต้องการของผู้ใช้งาน	4.66	0.50	ดีมาก
สรุปผลการประเมิน	4.60	0.50	ดีมาก

การประเมินผลการดำเนินงานใช้การประเมินความพึงพอใจ โดยกลุ่มผู้ใช้ คือ นักศึกษาระดับปริญญาตรีชั้นปีที่ 1 จำนวน 35 คน และชั้นปีที่ 2 จำนวน 35 คน รวม 70 คน โดยมีเกณฑ์การให้คะแนนแบบ 5 คะแนน ดังแสดงในตารางที่ 2

ตารางที่ 2: ผลการประเมินความพึงพอใจของผู้ใช้

รายการประเมิน	ระดับคุณภาพ		
	X	S.D.	เชิงคุณภาพ
1. ด้านความสะดวกและง่ายต่อ การใช้งานสื่อประกอบการสอน	4.77	0.42	ดีมาก
2. ด้านความช่วยเหลือเพื่อเสริมทักษะในการฝึกปฏิบัติการใช้งานโปรแกรมสำเร็จรูปในสำนักงาน	4.74	0.44	ดีมาก
3. ด้านความปลอดภัยของระบบในการเข้าใช้สื่อประกอบการสอน	4.62	0.49	ดีมาก
4. ด้านผลการประเมินผลความเข้าใจหลังเรียนผ่านสื่อประกอบการสอน	4.74	0.44	ดีมาก
สรุปผลการประเมิน	4.72	0.45	ดีมาก

5. สรุป

จากปัญหาด้านการฝึกปฏิบัติการใช้งานโปรแกรมสำเร็จรูป ทางผู้วิจัยได้ทำการแก้ไขปัญหาโดยการพัฒนาสื่อประกอบการสอนแบบมีส่วนร่วมออนไลน์ด้วยการเข้าถึงข้อมูลผ่านบาร์โค้ด 2 มิติเพื่อวัดความสนใจในเนื้อหาของผู้เรียนและเสริมทักษะการฝึกปฏิบัติใช้งาน โปรแกรมสำเร็จรูป ซึ่งผ่านการประเมินคุณภาพของสื่อประกอบการสอนและระบบวัดความสนใจในเนื้อหาของผู้เรียนจากผู้เชี่ยวชาญด้านการพัฒนาสื่อประกอบการสอนและระบบสารสนเทศจำนวน 2 ท่าน ได้ผลอยู่ในระดับดีมากโดยมีค่าเฉลี่ยอยู่ที่ 4.60 และค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.50 จากนั้นได้ทำการทดลองให้ผู้เรียนเป็นผู้ใช้งานสื่อประกอบการสอนผ่านบาร์โค้ด 2 มิติ ซึ่งผลการประเมินความพึงพอใจจากผู้ใช้งานจำนวน 70 คน อยู่ในระดับดีมากโดยมีค่าเฉลี่ยอยู่ที่ 4.72 และค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.45

เอกสารอ้างอิง

- [1] มณฑิธร กำบัง. แบบเรียนฝึกความสามารถผ่านเครือข่าย อินเทอร์เน็ตเรื่องการตัดต่อเสียงสำหรับการสร้างสื่อประกอบการ สอนวิทยาศาสตร์, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหาร ลาดกระบัง. สำนักหอสมุดกลาง, 2553.
- [2] อรนุช พันโท; อัยรดา พรเจริญการ. ศึกษาผลสัมฤทธิ์ของการสอน แบบในชั้นเรียนกับการสอนโดยใช้สื่อการสอนผ่านเครือข่าย คอมพิวเตอร์ กรณีศึกษา รายวิชาการเขียนโปรแกรมคอมพิวเตอร์ 1, มหาวิทยาลัยราชภัฏอุบลราชธานี.สำนักวิทยบริการและเทคโนโลยี สารสนเทศ, 2551
- [3] อภินันท์ จิตรเจริญ, การผลิตสื่อการสอนประกอบรายวิชา คอมพิวเตอร์และเทคโนโลยีสารสนเทศพื้นฐาน สำหรับนักเรียนชั้น มัธยมศึกษา, มหาวิทยาลัยเชียงใหม่. สำนักหอสมุด, 2548.
- [4] อัจฉรา เซขิงวิทย์ ; อธิป เกตุศิริ; สุกัญญา มิ่งมุก; อ้อมจิต กุติ, การพัฒนาสื่อการสอนบนเว็บไซต์เรื่องการสื่อสารข้อมูลและ เครือข่ายคอมพิวเตอร์, มหาวิทยาลัยราชภัฏอุบลราชธานี. สำนัก วิทยบริการและเทคโนโลยีสารสนเทศ, 2553
- [5] วิวัฒน์ชัย นิลวรรณ. 2546. การพัฒนาซอฟต์แวร์ระบบฐานข้อมูล บุคคลากรและนักเรียนในโรงเรียนประถมศึกษาโดยใช้บาร์โค้ด, สืบค้นเมื่อวันที่ 12 พฤศจิกายน 2556.
- [6] ศักรินทร์ ชะนะ, การประยุกต์ใช้คิวอาร์โค้ด ระบบระบุตำแหน่ง ด้วยเครือข่ายโทรศัพท์มือถือ และระบบฐานข้อมูลสำหรับ ผู้โดยสารอากาศยานขาเข้าที่โดยสารรถแท็กซี่สาธารณะสังกัดท่า อากาศยานสุวรรณภูมิ, มหาวิทยาลัยอีสเทิร์นเอเซีย. สำนัก หอสมุดกลาง, 2555.
- [7] เสษฐพงศ์ มะลิสวรรณ. 2555. QR Code. สืบค้นเมื่อวันที่ 12 พฤศจิกายน 2556. จาก<http://www.vcharkarn.com/varticle/41376>
- [8] อานาจิตวิทย์ หมูศิลป์, ระบบสารสนเทศเพื่อการบริหารครุภัณฑ์ คอมพิวเตอร์ด้วย Quick Response Code กรณีศึกษา สำนักงานการ ตรวจเงินแผ่นดิน, มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนคร เหนือ. สำนักหอสมุดกลาง, 2556. http://library.cmu.ac.th/digital_collection/eresearch/ [ออนไลน์]
- [9] ณัฐนันท์ ทรุชศิริ และ เศรษฐพัฒน์ โนทัย. 2556. การประยุกต์ใช้ บาร์โค้ดสองมิติชนิดคิวอาร์เพื่อการจัดการร้านค้า. สืบค้นเมื่อวันที่ 12 พฤศจิกายน 2556. จาก<http://www.mfu.ac.th/division/graduate/NGRC29th/PDF/PDF-ST/Oral/391.pdf>
- [10] ภกวัตรภักศรี. 2553. เครือข่ายสังคม (Social Networking). สืบค้นเมื่อ วันที่ 12 พฤศจิกายน 2556. จาก http://www.coe.or.th/e_engineers/knc_detail.php?id=170
- [11] Y.Swetake. 2013. How to create QRcode. สืบค้นเมื่อวันที่ 12 พฤศจิกายน 2556. จาก http://www.swetake.com/qrcode/qr1_en.html

การวิเคราะห์พฤติกรรมสำหรับการเลือกสมัครสาขาวิชาเรียนและการเปรียบเทียบตัวแบบ

พยากรณ์จำนวนนักศึกษาใหม่โดยใช้เทคนิคการทำเหมืองข้อมูล

The Behavior Analysis on the Applying Major Selection and the Comparison of Model to Forecast the Numbers of New Students Using Data Mining Technique

ธีรพงษ์ สังข์ศรี (Teerapong Sungsi)

โปรแกรมวิชาวิทยาการสารสนเทศ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครราชสีมา
evan.s.nrru@hotmail.com

บทคัดย่อ

ในปัจจุบันการรับสมัครสอบคัดเลือกนักศึกษาใหม่ของแต่ละสถาบันการศึกษาในระดับอุดมศึกษาได้มีรูปแบบเป็นนโยบายเชิงรุกมากขึ้นเช่น การประชาสัมพันธ์หลักสูตร การไปแนะแนวหลักสูตรตามโรงเรียนต่างๆ เป็นต้น เนื่องด้วยสถานการณ์แข่งขันที่สูงขึ้นของแต่ละสถาบัน งานวิจัยชิ้นนี้ได้นำแนวความคิดการทำเหมืองข้อมูล (Data Mining) มาช่วยในการวิเคราะห์ข้อมูลผู้สมัครเข้าศึกษาต่อซึ่งถูกจัดเก็บไว้เป็นฐานข้อมูลเพื่อช่วยในการวางแผนการรับสมัครนักศึกษาในอนาคต โดยสามารถแบ่งส่วนการทำงานออกได้เป็น 2 ส่วนหลักประกอบด้วย 1) การวิเคราะห์พฤติกรรมสำหรับการเลือกสมัครสาขาวิชาเรียน เป็นการประยุกต์ใช้เทคนิคการจัดกลุ่มข้อมูล (Clustering) แบบ Simple K-means สามารถแบ่งข้อมูลพฤติกรรมของผู้สมัครได้เป็น 4 กลุ่มและใช้การหากฎความสัมพันธ์ของข้อมูล (Association Rules) ด้วยเทคนิคอะพริออริ (Apriori) เพื่อหาความสัมพันธ์ของข้อมูลในแต่ละกลุ่มพฤติกรรมผู้สมัคร โดยใช้ค่าความเชื่อมั่นเท่ากับ 0.9 และ 2) การเปรียบเทียบตัวแบบพยากรณ์จำนวนนักศึกษาใหม่โดยตัวแบบพยากรณ์ที่ถูกสร้างขึ้นด้วยเทคนิคต้นไม้ช่วยตัดสินใจ (Decision Tree) มีค่าความถูกต้องเท่ากับ 93.76% และตัวแบบที่ถูกสร้างขึ้นด้วยเทคนิคโครงข่ายประสาทเทียม (Artificial Neural Network) แบบเพอร์เซ็ปตรอนหลายชั้น (Multi-layer Perceptron) มีค่าความถูกต้องเท่ากับ 93.60%

คำสำคัญ: การทำเหมืองข้อมูล ตัวแบบ พยากรณ์

Abstract

Nowadays, a recruitment examination of a new student for each institution of higher education has a more aggressive policy such as course public relations or Guidance to the school curriculums. Due to the higher competition of each institution. This research has gotten the data mining concept for helping the analysis of the candidate collected in the database and planning in the future. This research has two main components. 1) the behavior analysis on the applying major selection is to apply a clustering as K-means and Association Rules as Apriori for clustering into 4 groups and finding the association rule between the admission dataset with the minimum confidence as 0.9. 2) the comparison of forecasting model of the numbers of new students is to compare two models. The first model is constructed by a decision tree technique with an accuracy as 93.76%. And the second model is constructed by an artificial neural network with an accuracy as 93.60%

Keywords: Data Mining, Association Rules, Cluster, Forecasting model.

1. บทนำ

สถาบันอุดมศึกษามีกระบวนการสำหรับการคัดเลือกนักศึกษาใหม่เพื่อเข้าศึกษาในระดับปริญญาตรีด้วยวิธีการสอบคัดเลือกเป็นหลัก จากสถานการณ์แข่งขันที่สูงขึ้นสำหรับความต้องการนักศึกษาเข้าใหม่ในแต่ละสถาบันทั้งในด้านของคุณภาพการเรียนและจำนวนยอดของนักศึกษาใหม่ ส่งผลให้รูปแบบการรับสมัครเป็นนโยบายเชิงรุกมากขึ้น โดย

กระบวนการประชาสัมพันธ์ถือได้ว่าเป็นกระบวนการที่มีความสำคัญเป็นอย่างแรกสำหรับการสอบคัดเลือกนักศึกษาใหม่ของแต่ละสถาบันอุดมศึกษา โดยเฉพาะวิธีการประชาสัมพันธ์ตามโรงเรียนจำเป็นต้องใช้งบประมาณสูงและอาจไม่ตรงกับกลุ่มเป้าหมายที่ต้องการของแต่ละสาขาวิชา เนื่องจากไม่ได้มีการนำข้อมูลของผู้สมัครเข้าศึกษาต่อในแต่ละปีมาทำการวิเคราะห์ถึงพฤติกรรมในการเลือกสมัครเรียนเพื่อให้ทราบถึงกลุ่มพฤติกรรมในการเลือกสมัครเรียน ดังนั้นหากได้มีการนำข้อมูลของผู้สมัครเข้าศึกษาต่อแต่ละปีการศึกษาซึ่งทางสถาบันอุดมศึกษาแต่ละแห่งได้มีการจัดเก็บเป็นระบบฐานข้อมูลมาวิเคราะห์ด้วยเทคนิคการทำเหมืองข้อมูลจะทำให้ได้สารสนเทศที่เป็นประโยชน์อย่างมากต่อสถาบันอุดมศึกษาในการวางแผนสำหรับการรับสมัครนักศึกษาใหม่

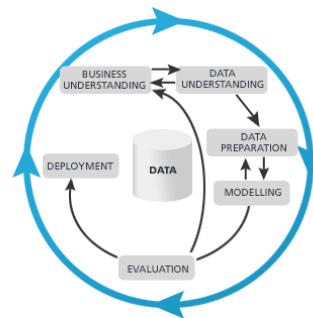
งานวิจัยชิ้นนี้มีวัตถุประสงค์เพื่อวิเคราะห์พฤติกรรมสำหรับการเลือกสมัครสาขาวิชาเรียนของผู้สมัครเข้าศึกษาต่อ โดยใช้เทคนิคการจัดกลุ่มข้อมูล (Clustering) และเทคนิคการหากฎความสัมพันธ์ของข้อมูล (Association Rule Mining: ARM) ในการจัดกลุ่มข้อมูลพฤติกรรมในการเลือกสมัครสาขาวิชาของผู้สมัครศึกษาต่อและค้นหากฎความสัมพันธ์ของแต่ละกลุ่มพฤติกรรมผู้สมัครเพื่อวิเคราะห์ปัจจัยที่ส่งผลต่อการเลือกสมัครเรียนในสาขาวิชาต่างๆ ในแต่ละกลุ่มพฤติกรรม และการเปรียบเทียบตัวแบบการพยากรณ์สำหรับทำนายจำนวนของนักศึกษาใหม่ ซึ่งมีการเปรียบเทียบประสิทธิภาพระหว่างตัวแบบพยากรณ์ที่ถูกพัฒนาขึ้นด้วยเทคนิคต้นไม้ช่วยตัดสินใจ (Decision Tree) กับเทคนิคโครงข่ายประสาทเทียม (Artificial Neural Network: ANN) โดยเลือกศึกษาข้อมูลของผู้สมัครเข้าศึกษาต่อ ประจำปีการศึกษา 2552-2555 มหาวิทยาลัยราชภัฏนครราชสีมา เป็นกรณีศึกษา

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การทำเหมืองข้อมูล (Data Mining) [1]

การทำเหมืองข้อมูล (Data Mining) หมายถึงกระบวนการที่ใช้สำหรับสกัดหรือสืบค้นความรู้ที่น่าสนใจจากข้อมูลปริมาณมากๆ หรือเรียกว่า “Knowledge Discovery from very large Databases: KDD” โดยขั้นตอนการทำเหมืองข้อมูล สามารถแบ่งขั้นตอนการทำงานตามโมเดลมาตรฐาน

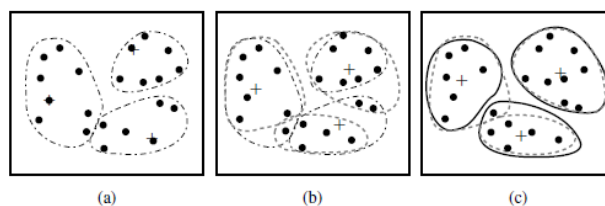
CRISP-DM (Cross-Industry Standard Process for Data Mining) ดังภาพที่ 1



ภาพที่ 1: โมเดลมาตรฐานการทำเหมืองข้อมูล CRISP-DM [1]

2.2 การจัดกลุ่มข้อมูล (Clustering) [2][3]

การจัดกลุ่มข้อมูลเป็นเทคนิคที่ใช้สำหรับการแบ่งกลุ่มข้อมูลออกเป็นกลุ่มย่อยตั้งแต่ 2 กลุ่มขึ้นไป โดยข้อมูลที่ถูกรวบรวมเป็นกลุ่มจะมีลักษณะของข้อมูลที่คล้ายคลึงกัน เพื่อสามารถวิเคราะห์ลักษณะเด่นของข้อมูลในแต่ละกลุ่มได้ เทคนิคพื้นฐานที่นิยมใช้สำหรับการจัดกลุ่มข้อมูลคือ K-means ซึ่งเป็นเทคนิคการเรียนรู้แบบไม่ต้องมีผู้สอน (Unsupervised Learning) จะเริ่มด้วยการแบ่งข้อมูลออกเป็น K กลุ่ม โดยแทนแต่ละกลุ่มด้วยค่าเฉลี่ยของกลุ่มสำหรับเป็นจุดศูนย์กลาง (centroid) ของกลุ่ม และทำการสร้างข้อมูลที่มีความสัมพันธ์ใกล้เคียงกับจุดศูนย์กลางมากที่สุดและคิดคำนวณจุดศูนย์กลางใหม่ ทำงานไปเรื่อยๆ จนกว่าค่าจุดศูนย์กลางจะไม่เปลี่ยนแปลง ดังภาพที่ 2



ภาพที่ 2: การจัดกลุ่มข้อมูลด้วยเทคนิค K-means [2]

จากภาพที่ 2 แสดงขั้นตอนการจัดกลุ่มข้อมูลที่ถูกกำหนดให้กลุ่มข้อมูลมีจำนวนเท่ากับ 3 กลุ่ม โดยมีลักษณะการสร้างกลุ่มข้อมูลดังภาพ (a) (b) และ (c) ตามลำดับ

2.3 การหากฎความสัมพันธ์ (Association Rules) [2][3]

การค้นหากฎความสัมพันธ์ของข้อมูลเป็นการค้นหารูปแบบความสัมพันธ์ที่เกิดขึ้นร่วมกันในรูปแบบของความถี่ในการเกิดความสัมพันธ์ที่ถูกเรียกว่า “Frequent Itemsets” และมีรูปแบบทั่วไปของกฎความสัมพันธ์อยู่ในลักษณะ $A \Rightarrow B$ โดยที่ A และ B คือชุดของข้อมูลที่เกิดขึ้นในรายการข้อมูล

(Itemsets) สำหรับการวัดค่าความสำคัญของกฎที่ถูกค้นหาได้ในชุดข้อมูลจะใช้ค่าความเชื่อมั่น (Confidence) และค่าสนับสนุน (Support) ดังคำนวณได้จากสมการที่ 1 และสมการที่ 2

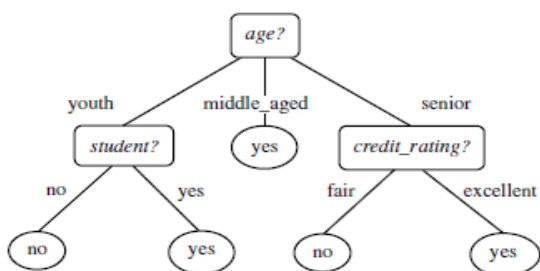
$$\text{support}(A \Rightarrow B) = P(A \cup B) = \frac{|A \cup B|}{|All|} \quad (1)$$

$$\text{confidence}(A \Rightarrow B) = P(B|A) = \frac{|A \cap B|}{|A|} \quad (2)$$

จากสมการค่าสนับสนุนคิดได้จากจำนวนความถี่สำหรับไอเทมเซต A และ B ที่เกิดขึ้นร่วมกันเทียบกับจำนวนรายการข้อมูลทั้งหมด และสำหรับค่าความเชื่อมั่นคิดได้จากจำนวนความถี่สำหรับไอเทมเซต A และ B ที่เกิดขึ้นร่วมกันเทียบกับจำนวนความถี่ที่ไอเทมเซต A เกิดขึ้นเพียงอย่างเดียว

2.4 การจำแนกชนิดข้อมูล (Classification) [2][3]

การจำแนกชนิดข้อมูลเป็นเทคนิคสำหรับการเรียนรู้แบบจำเป็นต้องมีผู้สอน (Supervised Learning) ซึ่งลักษณะการทำงานจะมีการแบ่งชุดของข้อมูลออกเป็น 2 ชุด โดยชุดข้อมูลส่วนแรกเรียกว่า "ชุดข้อมูลสำหรับการฝึกสอน" (Training set) เป็นชุดข้อมูลที่เอาไว้สำหรับให้เครื่องจักรเรียนรู้และสร้างตัวแบบการจำแนกข้อมูลขึ้นมา และชุดข้อมูลส่วนที่สองเรียกว่า "ชุดข้อมูลทดสอบ" (Testing set) เป็นชุดข้อมูลที่นำไปทดสอบประสิทธิภาพของตัวแบบการจำแนกชนิดข้อมูลที่ถูกสร้างขึ้นมา โดยเทคนิคพื้นฐานที่นิยมทั่วไปคือ เทคนิคต้นไม้ตัดสินใจ (Decision Tree) ดังภาพที่ 3 และโครงข่ายประสาทเทียม



ภาพที่ 3: การจำแนกชนิดข้อมูลด้วยต้นไม้ตัดสินใจ [2]

จากภาพที่ 3 แสดงโครงสร้างของต้นไม้ตัดสินใจ โดยลีฟ โหนด (Leaf node) นั้นจะเป็นชนิดข้อมูลที่ต้องการจำแนก (Class)

โครงสร้างของโครงข่ายประสาทเทียมแบบเพอร์เซปตรอนหลายชั้น (Multi-Layer Perceptron: MLP) [4] ซึ่งเป็นรูปแบบหนึ่งของโครงข่ายประสาทเทียมโดยใช้อัลกอริทึมในการผ่านค่าแบบย้อนกลับ (Backpropagation) เพื่อการส่งผ่านย้อนกลับในการปรับค่าน้ำหนักให้เหมาะสมในการแก้ไขข้อผิดพลาด (error correction)

2.5 งานวิจัยที่เกี่ยวข้อง

การวิเคราะห์พฤติกรรมสำหรับการเลือกสมัครสาขาวิชาเรียนและการเปรียบเทียบตัวแบบพยากรณ์จำนวนนักศึกษาใหม่โดยใช้การทำเหมืองข้อมูลได้ศึกษางานวิจัยที่เกี่ยวข้อง ดังนี้

งานวิจัยของจามรกุล และจิรารัตน์ [5] ได้สร้างตัวแบบพยากรณ์ลักษณะความเหมาะสมของนักศึกษาสาขาวิชาเทคโนโลยีสารสนเทศ โดยนำเทคนิคกฎการจำแนกเชิงความสัมพันธ์ (CARs: Class Association Rules) และแบ่งลักษณะความเหมาะสมของนักศึกษาตามระดับผลการเรียนคือ ระดับดีเยี่ยม ระดับดี ระดับปานกลาง และระดับต่ำ ผลการทดสอบตัวแบบพยากรณ์ที่สร้างขึ้นมีค่าความถูกต้องเฉลี่ยร้อยละ 83.81

งานวิจัยของชิดชนก ธนาวิวัฒน์และกฤษณะ [6] ได้นำเสนอการใช้เทคนิคต้นไม้ตัดสินใจเพื่อช่วยสำหรับการแก้ปัญหาการเลือกสาขาวิชาเรียนของนิสิตคณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ และแบ่งกลุ่มความเหมาะสมของนิสิตเป็น 2 กลุ่มคือ Good และ Bad มีค่าความถูกต้องร้อยละ 80

งานวิจัยของชลนิตา [7] ได้นำเทคนิคต้นไม้ตัดสินใจมาใช้สำหรับจำแนกสถานภาพการสำเร็จการศึกษาของนักศึกษา โดยแบ่งกลุ่มสถานภาพการสำเร็จการศึกษาออกเป็น 3 กลุ่มคือ สำเร็จการศึกษา ไม่สำเร็จการศึกษาและกำลังศึกษา โดยงานวิจัยดังกล่าวจะทำการเปรียบเทียบค่า precision และ recall ที่ได้กับตัวแบบของเทคนิคโครงข่ายประสาทเทียม

งานวิจัยของลาภ และมาลีรัตน์ [8] ได้นำเทคนิคต้นไม้ตัดสินใจมาใช้ในการสร้างตัวแบบการพยากรณ์สำหรับทำนายจำนวนนักศึกษาใหม่ โดยอัลกอริทึมสำหรับสร้างตัวแบบต้นไม้ตัดสินใจที่เลือกใช้คือ C4.5 ตัวแบบพยากรณ์ที่ได้มีค่าความถูกต้องมากกว่าร้อยละ 90

งานวิจัยของสุมาลย์ สิริภัทรและคำรณ [9] ได้ใช้โครงข่ายประสาทเทียมหลายชั้นแบบแบ็คพรอบพาเกชัน (Multilayer

Back propagation Neural Network) โดยมีวิธีการแบ่งข้อมูลเป็นอัตราส่วน 3 กลุ่มคือ 60:40 70:30 และ 80:20 ได้ค่าความถูกต้องมากที่สุดคือร้อยละ 87.98

3. วิธีการดำเนินงาน

งานวิจัยนี้ได้ใช้เครื่องมือในการทำเหมืองข้อมูลคือโปรแกรม WEKA 3.6.2 สำหรับดำเนินการวิจัยและมีวิธีดำเนินการวิจัยออกเป็นขั้นตอน ดังต่อไปนี้

3.1 การจัดเตรียมข้อมูล

งานวิจัยนี้ได้มีการจัดเตรียมข้อมูลโดยใช้ข้อมูลของผู้สมัครเข้าศึกษาต่อ มหาวิทยาลัยราชภัฏนครราชสีมา เฉพาะภาคปกติ จำนวนทั้งสิ้น 70,000 ข้อมูลผู้สมัครตั้งแต่ปีการศึกษา 2552-2555 ข้อมูลถูกปรับปรุงในส่วนของค่าข้อมูลที่ขาดหาย (Missing value) ได้มีการใช้วิธีการตัดเรคอร์ดที่ข้อมูลไม่สมบูรณ์ทิ้งทั้งหมด การเลือกแอตทริบิวต์เฉพาะที่เกี่ยวข้องและมีการปรับเปลี่ยนรูปแบบของข้อมูลบางแอตทริบิวต์เพื่อลดขนาดข้อมูลที่ใช้ในการประมวลผล โดยข้อมูลที่สมบูรณ์พร้อมสำหรับการนำไปวิเคราะห์ในงานวิจัยนี้ประกอบด้วย 7 แอตทริบิวต์ และมีจำนวนเรคอร์ดข้อมูล 42,000 เรคอร์ด สามารถแสดงตัวอย่างข้อมูลที่สมบูรณ์ ดังตารางที่ 1

ตารางที่ 1: ตัวอย่างข้อมูลที่ผ่านการจัดเตรียมข้อมูลอย่างสมบูรณ์

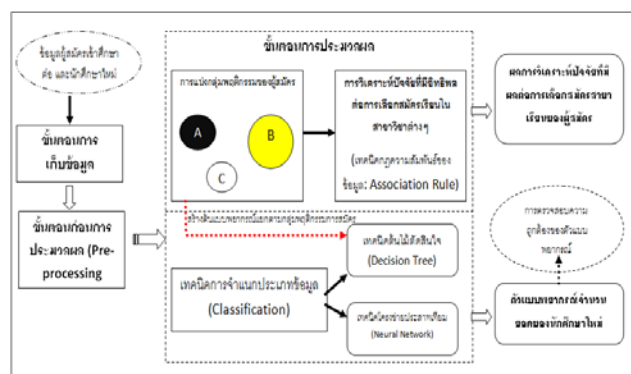
Sex	Grade	Type	School	Major1	Major2	Major3
F	Mid_G	Exam	OP	BM	BM	MS
M	Mid_G	Exam	AP	BM	BM	MS
M	High_G	Exam	AP	SC	ED	MS
...

จากตารางที่ 1 แสดงตารางข้อมูลผู้สมัครศึกษาต่อที่ผ่านการปรับปรุงสมบูรณ์แล้ว โดยประกอบด้วยแอตทริบิวต์ Sex (เพศ) Grade (เกรดเฉลี่ยแบ่งเป็น High_G ระดับคะแนนมากกว่า 3.00 Mid_G ระดับคะแนน 2.00-3.00 และ Low_G ระดับคะแนนน้อยกว่า 2.00) Type (ประเภทการคัดเลือกแบ่งเป็น Exam คือประเภทสอบคัดเลือก Act คือประเภทกิจกรรม Best คือประเภททุนเพชรราชภัฏ และ Qua คือประเภทโควต้า) School (ประเภท

โรงเรียนแบ่งเป็น OP คือโรงเรียนต่างจังหวัด OM คือโรงเรียนต่างอำเภอ และ AM คือโรงเรียนในอำเภอเมืองนครราชสีมา) Major1 (สาขาวิชาคณะที่เลือกสมัครอันดับหนึ่ง) Major2 (สาขาวิชาคณะที่เลือกสมัครอันดับสอง) Major3 (สาขาวิชาคณะที่เลือกสมัครอันดับสาม)

3.2 การวิเคราะห์พฤติกรรมสำหรับการเลือกสมัครสาขาวิชาเรียนและเปรียบเทียบตัวแบบพยากรณ์จำนวนยอดนักศึกษาใหม่

งานวิจัยนี้ได้แบ่งส่วนงานออกเป็น 2 ส่วนหลักคือ 1) การวิเคราะห์พฤติกรรมสำหรับการเลือกสมัครสาขาวิชาเรียน และ 2) การเปรียบเทียบตัวแบบพยากรณ์จำนวนยอดนักศึกษาใหม่ สามารถแสดงการทำงาน ดังภาพที่ 4



ภาพที่ 4: กรอบแนวคิดการทำวิจัย

3.2.1 การวิเคราะห์พฤติกรรมในการเลือกสมัครสาขาวิชาเรียน โดยขั้นแรกจะนำข้อมูลที่จัดเตรียมไว้มาทำการจัดกลุ่มข้อมูล (Clustering) ด้วยเทคนิคพื้นฐานทั่วไปคือ Simple K-means ด้วยโปรแกรม WEKA สามารถจัดกลุ่มพฤติกรรมสำหรับผู้สมัครเรียนที่เหมาะสมได้เป็น 4 กลุ่มประกอบด้วย

กลุ่มที่หนึ่ง ผู้สมัครมีเกรดเฉลี่ยอยู่ระดับปานกลาง และเรียนอยู่ในโรงเรียนนอกอำเภอเมือง

กลุ่มที่สอง ผู้สมัครมีเกรดเฉลี่ยอยู่ระดับสูง และเรียนอยู่ในโรงเรียนนอกอำเภอเมือง

กลุ่มที่สาม ผู้สมัครมีเกรดเฉลี่ยอยู่ระดับสูง และเรียนอยู่ในโรงเรียนต่างจังหวัด

กลุ่มที่สี่ ผู้สมัครมีเกรดเฉลี่ยอยู่ระดับปานกลาง และเรียนอยู่ในโรงเรียนต่างจังหวัด

จากนั้น จะทำการหาความสัมพันธ์ของข้อมูล (Association Rules) ของแต่ละกลุ่มพฤติกรรมของผู้สมัครเข้า

เป็นการสอบคัดเลือก จะมีการเลือกสาขาอันดับ2 เป็นคณะมนุษยศาสตร์และสังคมศาสตร์

4.2 การเปรียบเทียบตัวแบบพยากรณ์จำนวนนักศึกษาใหม่

การวัดประสิทธิภาพของตัวแบบพยากรณ์จำนวนนักศึกษาใหม่ด้วยเทคนิคต้นไม้ตัดสินใจและโครงข่ายประสาทเทียมจะใช้ค่าความถูกต้อง (Accuracy) ในการทำนายผล โดยมีการทดสอบตัวแบบพยากรณ์ด้วยวิธีการแบ่งชุดข้อมูลเท่ากับ 70:30 (ข้อมูลการฝึกสอนร้อยละ 70 และข้อมูลการทดสอบตัวแบบร้อยละ 30) สามารถแสดงการเปรียบเทียบได้ ดังตารางที่ 2

ตารางที่ 2: ผลการเปรียบเทียบประสิทธิภาพตัวแบบพยากรณ์จำนวนนักศึกษาใหม่

	ต้นไม้ตัดสินใจ	โครงข่ายประสาทเทียม
ความถูกต้อง	ร้อยละ 93.76	ร้อยละ 93.60

จากตารางที่ 2 แสดงให้เห็นว่าตัวแบบพยากรณ์จำนวนยอดของนักศึกษาใหม่ที่ถูกร่างขึ้นด้วยเทคนิคต้นไม้ตัดสินใจมีความถูกต้องเท่ากับ 93.76% ซึ่งสูงกว่าตัวแบบพยากรณ์ที่ถูกร่างขึ้นด้วยโครงข่ายประสาทเทียมเพียงเล็กน้อย

5. สรุป

งานวิจัยนี้เป็นการประยุกต์ใช้เทคนิคการทำเหมืองข้อมูลเพื่อนำมาวิเคราะห์พฤติกรรมกรรมการเลือกสมัครสาขาวิชาเรียนของผู้สมัครเข้าศึกษาต่อในระดับปริญญาตรี มหาวิทยาลัยราชภัฏนครราชสีมา และนำเสนอการเปรียบเทียบประสิทธิภาพของตัวแบบพยากรณ์จำนวนนักศึกษาใหม่ด้วยเทคนิคต้นไม้ตัดสินใจและโครงข่ายประสาทเทียม โดยมีผลการทดลองดังนี้ กลุ่มผู้สมัครที่สังกัดโรงเรียนนอกอำเภอเมืองจะมีการเลือกสมัครในสาขาวิชาของคณะวิทยาศาสตร์และครุศาสตร์ โดยเฉพาะถ้าผู้สมัครเป็นเพศหญิงจะเลือกสมัครสาขาวิชาเรียนในคณะครุศาสตร์ หรือถ้าผู้สมัครเรียนอยู่ในโรงเรียนต่างจังหวัด จะเลือกสมัครสาขาของคณะวิทยาศาสตร์และเทคโนโลยี เป็นต้น และตัวแบบพยากรณ์จำนวนนักศึกษาใหม่ที่ถูกร่างขึ้นด้วยเทคนิคต้นไม้ตัดสินใจมีความถูกต้องสูงกว่าตัวแบบพยากรณ์ที่ถูกร่างขึ้นด้วยเทคนิคโครงข่ายประสาทเทียมเพียง 0.16% เท่านั้น

จากผลการทดลองแสดงให้เห็นว่ากฎความสัมพันธ์ที่ค้นหาได้จากข้อมูลผู้สมัครเข้าศึกษาต่อ นั้นสามารถบ่งบอกพฤติกรรม

การเลือกสมัครสาขาวิชาเรียนได้เพียงวงแคบ และตัวแบบพยากรณ์จำนวนนักศึกษายังเป็นเพียงการทำนายแต่ละข้อมูล ผู้สมัครเท่านั้นจึงต้องมีการสรุปจำนวนยอดนักศึกษาที่ทำนายได้อีกครั้ง

เอกสารอ้างอิง

- [1] Chapman Pete, Clinton Julian, Kerber Randy, eds. CRISP-DM 1.0 Step-by-step data mining guide. SPSS Inc., USA. 2000.
- [2] Jaiwei Han and Micheline Kamber. Data Mining Concepts and Techniques, Second Edition. University of Illinois at Urbana-Champaign. 2006.
- [3] S. Prabhu and N. Venkatesan. Data mining and Warehousing. New Age International Publicshers. 2007.
- [4] เอกรินทร์ แซ่เอ็ง. โครงข่ายประสาทเทียมกับการประยุกต์ใช้งาน. แผนกสารสนเทศ สำนักวิชาการ วิทยาลัยนอร์ท, กรุงเทพ. [ออนไลน์] http://ba.northbkk.ac.th/download/File_PDF/neural_network1.pdf
- [5] จามรกุล เหล่าเกียรติคุณ และจิรารัตน์ สิทธิวรชาติ. ตัวแบบพยากรณ์ลักษณะความเหมาะสมของนักศึกษาใหม่สาขาวิชาเทคโนโลยีสารสนเทศด้วยกฎการจำแนกประเภทเชิงความสัมพันธ์. วารสารเทคโนโลยีสารสนเทศ. ปีที่ 4. ฉบับที่ 8. พ.ศ. 2551 หน้าที่ 7-13.
- [6] ชิดชนก ส่งศิริ, ธนาวิทย์ รักธรรมานนท์ และกฤษณะ ไวยมัย. การใช้เทคนิค Data Mining เพื่อค้นหาภาควิชาที่เหมาะสมที่สุดให้กับนิสิต. การประชุมทางวิชาการของมหาวิทยาลัยเกษตรศาสตร์ ครั้งที่ 39. พ.ศ. 2544 หน้าที่ 43-50.
- [7] ชลนิศา สาระ. การจำแนกกลุ่มสถานภาพการสำเร็จการศึกษาโดยแบบจำลองต้นไม้ตัดสินใจ. ภาควิชาวิทยาการคอมพิวเตอร์และสารสนเทศ บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ พ.ศ. 2550.
- [8] ลาก พุ่มศิริรัฐ และมาลีรัตน์ โสคานิล. การพัฒนาตัวแบบการพยากรณ์จำนวนนักศึกษาใหม่โดยใช้กฎการจำแนกเทคนิคต้นไม้ตัดสินใจ. วารสารวิชาการและวิจัย มทรพระนคร ฉบับพิเศษ. ราช มงคลพระนครวิชาการ ครั้งที่ 2 พ.ศ. 2554.
- [9] สุมาลัย นุชิต, สิริภัทร เชื้อชาชาญวัฒนา และคำรณ สุนดี. การวิเคราะห์ปัจจัยที่มีอิทธิพลต่อการเพิ่มผลของจำนวนนักศึกษาระดับศึกษา: วิทยาลัยการอาชีพวังสะพุง. The National Conference on Computer Information Technologies, 2010.

การสร้างภาพเสมือนสารสนเทศของนักเรียนโรงเรียน ในสังกัดองค์การบริหารส่วนจังหวัดนครราชสีมา The Visualization of Student Information in the Schools of Nakhon Ratchasima Provincial Administrative Organization

นิคม วุฒิยา (NikhomWuttiya)¹, คำรณ สุณัติ (KhamronSunat)² และสิริภัทร เชี่ยวชาญวัฒนา (SirapatChiewchanwattana)³

^{1,2,3}ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น
nikhom@kkumail.com¹, khamron_sunat@yahoo.com², sunkra@kku.ac.th³

บทคัดย่อ

การจัดสรรงบประมาณให้แก่สถานศึกษาในแต่ละปีนั้นมักจะสร้างปัญหาให้แก่หน่วยงานต้นสังกัดเป็นอย่างมาก เนื่องจากหน่วยงานต้นสังกัดไม่ทราบข้อมูลพื้นฐานโดยรวมของแต่ละโรงเรียนที่จะนำมาวิเคราะห์ปัจจัยต่างๆในการพัฒนาสถานศึกษาให้เท่าเทียมกันและตรงกับความต้องการของแต่ละโรงเรียน ซึ่งปัจจุบันโรงเรียนสังกัดกรมส่งเสริมปกครองท้องถิ่นได้ให้มีการจัดเก็บข้อมูลของสถานศึกษาในระบบข้อมูลสารสนเทศทางการศึกษาท้องถิ่น (SIS) แต่ข้อมูลสารสนเทศเหล่านั้นมีความหลากหลาย ซึ่งยากต่อการนำมาวิเคราะห์และวางแผนในการจัดสรรงบประมาณให้กับสถานศึกษาและตรงกับบริบทพร้อมทั้งตรงกับความต้องการของสถานศึกษา

งานวิจัยนี้เป็นการสร้างภาพเสมือนสารสนเทศของนักเรียนโรงเรียนในสังกัดองค์การบริหารส่วนจังหวัดนครราชสีมา โดยใช้กลุ่มตัวอย่างจำนวน 57 แห่ง ซึ่งใช้เทคนิคการแบ่งกลุ่มแบบแผนผังการจัดระเบียบตัวเอง ที่มีชุดข้อมูลสารสนเทศทางการศึกษาท้องถิ่น 3 ด้าน ได้แก่ ด้านจำนวนนักเรียน ด้านข้อมูลการเจริญเติบโตของนักเรียน และด้านรายได้ผู้ปกครองนักเรียน ที่มีผลต่อการจัดสรรงบประมาณประจำปี จากการทดลองพบว่า ข้อมูลด้านจำนวนนักเรียนมีความสัมพันธ์ในเชิงจัดสรรงบประมาณให้แก่สถานศึกษา ข้อมูลด้านการเจริญเติบโตของนักเรียนมีความสัมพันธ์ต่อส่วนสูงกับน้ำหนัก และข้อมูลด้านรายได้ผู้ปกครองมีความสัมพันธ์กับความยากจนของนักเรียน ซึ่งข้อมูลทั้ง 3 ด้านนี้สามารถนำไปวางแผนการช่วยเหลือด้านงบประมาณให้แก่สถานศึกษาได้

คำสำคัญ: แผนผังการจัดระเบียบตัวเอง สารสนเทศทางการศึกษาท้องถิ่น ระบบข้อมูลสารสนเทศทางการศึกษาท้องถิ่น

Abstract

Lack of basic information of schools causes difficulties in managing the budget to equally meet need of each school by the headquarter offices. The Department of Local Administration has collected student and schools information via School Information System (SIS). The collected data, however, was diverse and that hard to employ as database for and financial planning and budget distribution to meet the need of institutions.

This research aims to provide The visualization of student Information using 57 sample schools in NakhonRatchasima under Provincial Administrative Organization. Own group organized technique with three groups of education information, namely the number of students, the growth data of the students, and parent income.

The results show that all three information data are correlated and can be used to assist annual budget planning.

Keyword: SOM, Information local school, Student Information Schools, SIS

1. บทนำ

สังคมไทยในปัจจุบันการพัฒนาชาตินั้นจำเป็นต้องอาศัยทรัพยากรมนุษย์ เป็นกำลังขับเคลื่อนกลไกทางสังคม เศรษฐกิจและวัฒนธรรม เพื่อให้เกิดความเจริญรุ่งเรืองของชาติ ซึ่งทรัพยากรมนุษย์นั้นจะมีคุณภาพได้ต้องอาศัยความรู้ความสามารถเฉพาะด้านของแต่ละบุคคล

องค์การบริหารส่วนจังหวัดนครราชสีมา มีโรงเรียนในสังกัดทั้งสิ้น 58 โรงเรียน ซึ่งล้วนแต่เป็นระดับมัธยม มีการจัดเก็บข้อมูลสารสนเทศทางการศึกษาท้องถิ่นของโรงเรียน ในระบบข้อมูลสารสนเทศทางการศึกษาท้องถิ่น (Student Information Schools : SIS) ที่สามารถนำข้อมูลมาพิจารณาจัดสรรงบประมาณเงินอุดหนุนเฉพาะกิจให้แก่ องค์การปกครองส่วนท้องถิ่นที่มีโรงเรียนในสังกัด

ปัจจุบันข้อมูลสารสนเทศทางการศึกษาท้องถิ่นขององค์การบริหารส่วนจังหวัดนครราชสีมา เป็นข้อมูลเชิงตัวเลขยากต่อการเข้าใจ จึงมีการนำข้อมูลเชิงตัวเลขนั้นมาสร้างภาพเสมือนใช้ในการวิเคราะห์หาปัจจัยการพัฒนาคุณภาพการศึกษาให้มีประสิทธิภาพมากขึ้น

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

เทคนิคการแบ่งกลุ่มแบบแผนผังการจัดระเบียบตัวเอง (Self-Organizing Map) การแบ่งกลุ่มแบบแผนผังการจัดระเบียบตัวเอง (Self- Organizing Map) [2] พัฒนาขึ้นในปี ค.ศ. 1982 โดยได้แนวคิดมาจากโครงข่ายประสาทเทียม (Artificial Neural Network) ซึ่งมีความสามารถในการเรียนรู้ที่คล้ายคลึงกับสมองของมนุษย์ โดยขั้นตอนการเรียนรู้นั้นอาศัยการนำเข้าข้อมูลเพื่อสร้างแบบจำลองสำหรับการพยากรณ์ของระบบงานต่างๆ วิธีแบบแผนผังการจัดระเบียบตัวเอง ของโคโฮเนนเป็นโครงข่ายประสาทเทียมแบบไม่มีชั้นซ่อน (Hidden layer) ดังนั้นจึงมีเพียง 2 ชั้นคือ ชั้นนำข้อมูลเข้า (Input Layer) และชั้นผลลัพธ์ (Output Layer)

แนวคิดการแบ่งกลุ่มแบบแผนผังการจัดระเบียบตัวเอง คือ การทำซ้ำของข้อมูลในแต่ละชุดของข้อมูล เพื่อหาค่าน้ำหนักของข้อมูลที่มีอยู่ทั้งหมดตามจำนวนกลุ่มที่ต้องการแบ่ง โดยสุ่ม

ค่าน้ำหนัก (W) ให้กับข้อมูล กำหนดให้ค่าความผิดพลาดที่ยอมรับได้ (α) อยู่ในช่วง 0 ถึง 1 จากนั้นเลือกข้อมูลนำเข้าเพื่อคำนวณหาระยะห่างของข้อมูลกับน้ำหนัก (D_j) ได้ดังสมการ

$$D_j = \sqrt{\sum_{i=1}^p (W_{ij} - X_i)^2} ; j = 1,2,3, \dots, k \quad (1)$$

โดยที่ W_{ij} คือ น้ำหนักของข้อมูลกลุ่มที่ j ตัวแปรที่ i
 X_i คือ ค่าของข้อมูลในตัวแปรที่ i
 D_j คือ ระยะทางระหว่างข้อมูลกับค่าน้ำหนัก
 p คือ จำนวนตัวแปรทั้งหมด

จากนั้นเปรียบเทียบค่าระยะห่างของข้อมูลกับน้ำหนัก (D_j) ในแต่ละกลุ่มข้อมูลว่าค่าของกลุ่มใดมีค่าน้อยที่สุดแล้วนำข้อมูลของกลุ่มที่มีค่าระยะห่างน้อยที่สุดมาคำนวณหาค่าน้ำหนักใหม่ดังสมการ

$$W_{ij}(\text{new}) = W_{ij}(\text{old}) + \alpha(X_i - W_{ij}(\text{old})) \quad (2)$$

โดยที่ X_i คือ ค่าของข้อมูลในตัวแปรที่ i
 $W_{ij}(\text{new})$ คือ น้ำหนักใหม่ของข้อมูลกลุ่มที่ j ตัวแปรที่ i
 $W_{ij}(\text{old})$ คือ น้ำหนักเก่าของข้อมูลกลุ่มที่ j ตัวแปรที่ i
 α คือ ค่าความผิดพลาดที่ยอมรับได้มีค่าอยู่ในช่วง 0 ถึง 1

จากนั้นปรับลดค่า α ทำซ้ำใหม่ จนกว่าจะไม่มีมีการเปลี่ยนแปลงของค่าน้ำหนัก (W)

ค่าน้ำหนักของข้อมูล (W) สามารถเขียนรูปเมทริกซ์ได้ดังนี้

$$W = \begin{bmatrix} w_1 & w_2 & \dots & w_k \\ w_{11} & w_{12} & \dots & w_{1k} \\ w_{21} & w_{22} & \dots & w_{2k} \\ \vdots & \vdots & \dots & \vdots \\ w_{p1} & w_{p2} & \dots & w_{pk} \end{bmatrix}_{p \times k} \quad (3)$$

เมื่อ p คือจำนวนตัวแปร และ k คือจำนวนกลุ่มซึ่ง $W_j = W_1, W_2, W_3, \dots, W_k$

B. Martín-del-Brío, A. Bono-Nuez, และ N. Medrano-Marqués [3] ได้ศึกษาเรื่อง Sale-Organizing Map for Embedded Microprocessor Selection ได้นำวิธีการแผนผังจัดระเบียบตนเองใช้ในการประมวลข้อมูลออกมาเป็นภาพ เพื่อใช้ในการวิเคราะห์คุณสมบัติของไมโครโปรเซสเซอร์ โดยมีองค์ประกอบ 6 ประเภทคือ ความเร็ว ระบายความร้อน การนำเข้าและระบายความร้อน ออก แรม กำลังไฟ และราคา เพื่อนำไปใช้ในการเลือกไมโครโปรเซสเซอร์ให้เหมาะสมกับลักษณะของงานซึ่งผลลัพธ์จะแสดงในรูปแบบของภาพของแต่ละคุณสมบัติตามความสามารถของไมโครโปรเซสเซอร์แต่ละตัวด้วยระดับสีที่แตกต่างกัน

Li และ F. Lin [7] ได้ศึกษาเรื่อง Customer segmentation analysis based on SOM clustering ซึ่งได้นำวิธีการแผนผังจัดระเบียบตนเองและวิธีเคมินและวิธีของการทำเหมืองข้อมูลมาจัดกลุ่มข้อมูลพฤติกรรมของลูกค้า

C.-S. Lee และ Y. P. Singh [4] ได้ศึกษาเรื่อง Student modeling using principal component analysis of SOM clusters โดยการสร้างแบบจำลองของนักเรียนโดยใช้วิธี machine-learning เพื่อศึกษาเปรียบเทียบเพื่อใช้ในการวิเคราะห์พร้อมจัดกลุ่ม โดยใช้วิธีการ SOM

อาทิษฐ์ กลีบรัง [2] ได้ศึกษาเรื่องการสร้างภาพเสมือนข้อมูลปัญหาเพื่อวิเคราะห์ความแตกต่างระหว่างผู้เรียน ได้นำวิธีการจัดลำดับ วิธีการแผนผังจัดระเบียบตนเอง และเทคนิคแผนผังการจัดระเบียบตนเองแบบเติบโตตามลำดับขั้น เพื่อหาเทคนิคที่เหมาะสมในวิเคราะห์ประสิทธิภาพผู้เรียนทั้งหมด 8 ด้านได้แก่ ด้านสติปัญญา ด้านภาษา ด้านตรรกะและคณิตศาสตร์ ด้านมิติสัมพันธ์ ด้านร่างกายและการเคลื่อนไหว ด้านดนตรี ด้านมนุษยสัมพันธ์ ด้านตนเองและการเข้าใจตนเอง และด้านการเข้าใจธรรมชาติ โดยแสดงผลออกมาในรูปแบบแผนผังแสดงระดับสติปัญญาของผู้เรียนในแต่ละด้านจากระดับสีขาว – ดำตามความโดดเด่นของสติปัญญา

นิรันดร์ สารรัมย์ [1] ได้ศึกษาเรื่องการจัดกลุ่มโรงเรียนคุณภาพตามมาตรฐานการศึกษาโดยใช้แผนผังการจัด

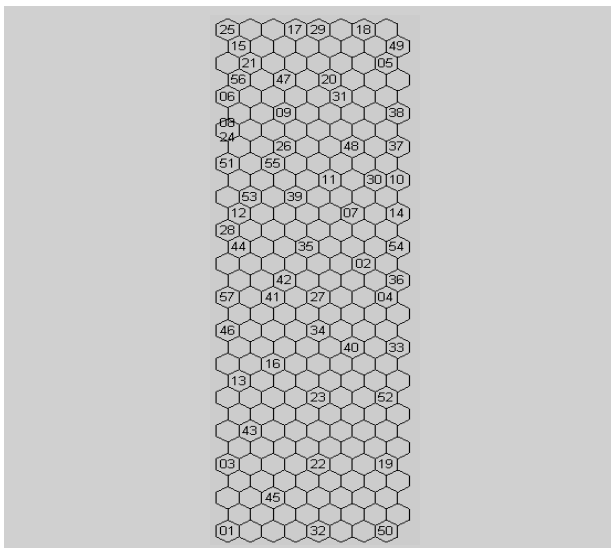
ระเบียบตัวเอง:กรณีศึกษาโรงเรียนสังกัดองค์การบริหารส่วนจังหวัดนครราชสีมา มีการจัดกลุ่มโรงเรียนคุณภาพตามมาตรฐานการศึกษาของกลุ่มตัวอย่าง 58 โรงเรียนโดยการใช้แผนผังการจัดระเบียบตัวเอง (SOM) เพื่อให้สถานศึกษาทราบปัญหาและจุดบกพร่องเมื่อเทียบกับสถานศึกษาอื่น โดยจะแสดงเป็นแผนภาพเสมือนข้อมูล (Data Visualization) ซึ่งจะบอกถึงคุณภาพของสถานศึกษาในแต่ละมาตรฐาน พร้อมทั้งบอกถึงกลุ่มของสถานศึกษาว่าอยู่ในกลุ่มใด โดยการนำเอาค่าเฉลี่ยของคะแนนแบบอิงเกณฑ์และอิงสถานศึกษาจากการประเมินของสำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์การมหาชน) รอบที่ 2 ซึ่งมี 14 ปีจ้อยจาก 3 ด้าน คือ ด้านผู้เรียน ด้านครูและด้านผู้บริหาร ผลการทดลองพบว่าโรงเรียนกลุ่มตัวอย่าง ส่วนใหญ่อยู่ในกลุ่มดีมีทั้งหมด 22 โรงเรียน และโรงเรียนที่อยู่ในกลุ่มปรับปรุงมี 6 โรงเรียน โดยภาพรวมของกลุ่มสถานศึกษาตัวอย่างทั้ง 58 โรงเรียน มาตรฐานที่ต้องปรับปรุงคือ มาตรฐานที่ 5 ผู้เรียนมีความรู้และทักษะที่จำเป็นตามหลักสูตร และมาตรฐานที่ 8 ครูมีคุณวุฒิความรู้/ความสามารถตรงกับงานที่รับผิดชอบและครูเพียงพอ

3. วิธีการดำเนินการวิจัย

มีการนำข้อมูลสารสนเทศทางการศึกษาท้องถิ่น ด้านจำนวนนักเรียน ด้านข้อมูลการเจริญเติบโตของนักเรียน ด้านรายได้ผู้ปกครองนักเรียน ที่สามารถจัดสรรงบประมาณเงินอุดหนุนเฉพาะกิจให้แก่โรงเรียนในสังกัด เช่น รายการครุภัณฑ์ Tablet รายการค่าปัจจัยพื้นฐานให้แก่เด็กนักเรียนยากจน รายการโครงการสนับสนุนการจัดการศึกษาโดยไม่เสียค่าใช้จ่าย 15 ปีจากนั้นนำข้อมูลมาวิเคราะห์เพื่อจัดกลุ่มการสร้างภาพเสมือนสารสนเทศของนักเรียนโรงเรียนในสังกัดองค์การบริหารส่วนจังหวัดนครราชสีมา จำนวน 57 โรงเรียน โดยใช้วิธีแผนผังการจัดระเบียบตัวเอง (Self- Organizing Map)

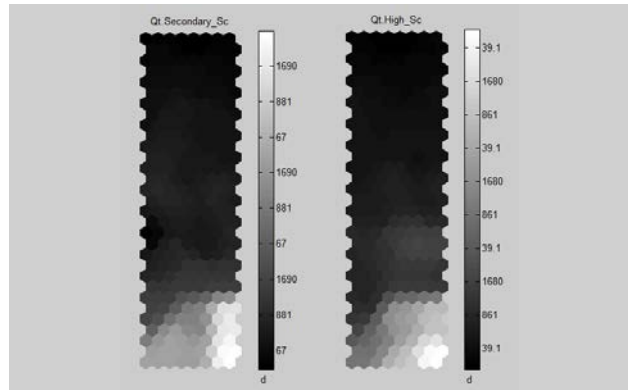
4. ผลการทดลอง

เมื่อนำชุดข้อมูลทดลองทั้ง 3 ด้าน คือ ด้านจำนวนนักเรียน ด้านข้อมูลการเจริญเติบโตของนักเรียน และด้านรายได้ผู้ปกครองนักเรียน จะได้ขอบเขตของโรงเรียนในสังกัดองค์การบริหารส่วนจังหวัดนครราชสีมา 57 แห่ง เป็นแผนผังข้อมูลที่ปรับระดับแล้วสร้างเป็นภาพเสมือนข้อมูลด้วยแผนผังจัดระเบียบตนเอง โดยโรงเรียนที่มีค่าโดยรวมสูงจะอยู่ด้านล่างของแผนผังและโรงเรียนที่มีค่าโดยรวมน้อยจะอยู่ด้านบนของแผนผัง ดังภาพที่ 1

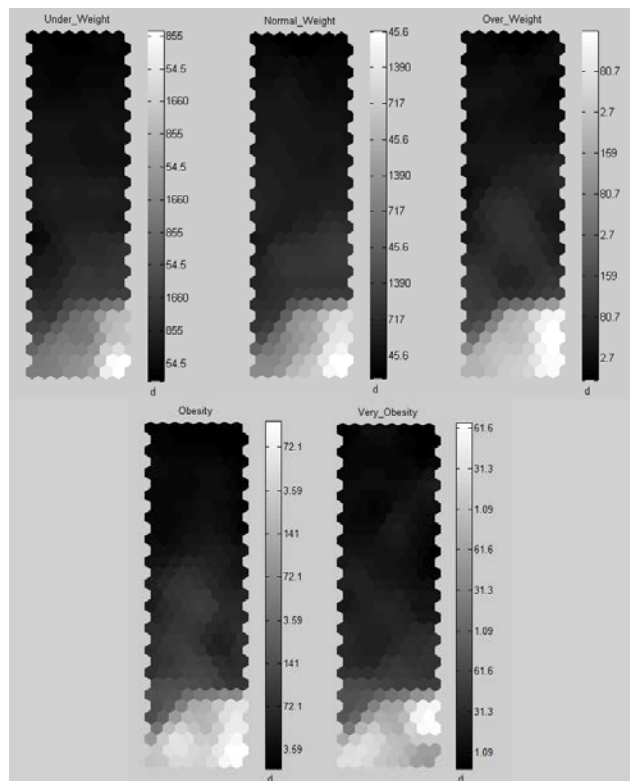


ภาพที่ 1: ขอบเขตของโรงเรียนในสังกัดองค์การบริหารส่วนจังหวัดนครราชสีมาทั้ง 57 แห่ง

ภาพที่ 2 จะเห็นได้ว่าแผนภาพด้านซ้ายเป็นระดับชั้นมัธยมศึกษาตอนต้น ไหลช่วงแถบสีจางคือโรงเรียนที่มีจำนวนนักเรียนมากและมีค่าคะแนนมาก ส่วนโรงเรียนที่มีจำนวนนักเรียนน้อยลงมาจะอยู่ในช่วงแถบสีเข้มและช่วงค่าคะแนนน้อย ถัดจากนั้นแผนภาพด้านขวาเป็นระดับชั้นมัธยมศึกษาตอนปลายที่มีจำนวนนักเรียนมากจะเป็นโรงเรียนที่อยู่ในช่วงแถบสีจางและมีค่าคะแนนมาก และโรงเรียนที่มีจำนวนนักเรียนระดับมัธยมศึกษาตอนปลายน้อยลงมาจะอยู่ในช่วงแถบสีเข้มและช่วงค่าคะแนนน้อยลงด้วย



ภาพที่ 2: ผลการทดลองด้านจำนวนนักเรียนระดับมัธยมศึกษาตอนต้นและนักเรียนระดับมัธยมศึกษาตอนปลาย

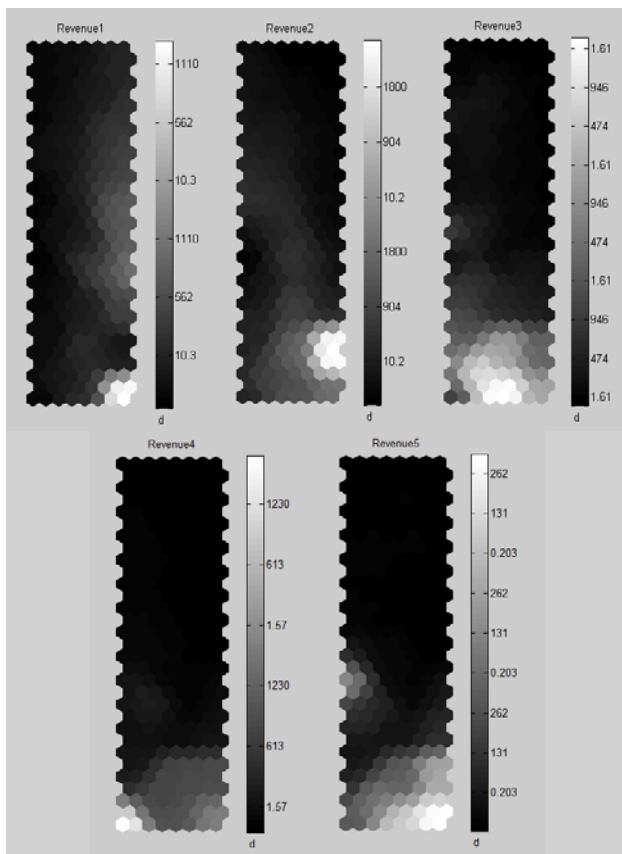


ภาพที่ 3: ผลการทดลองด้านการเจริญเติบโตของนักเรียน

จากภาพที่ 3 เป็นการทดลองด้านการเจริญเติบโตของนักเรียน (Body Mass Index : BMI) โดยภาพแรกจากซ้ายมือเป็นโรงเรียนที่มีนักเรียนน้ำหนักต่ำกว่าเกณฑ์ ค่า BMI น้อยกว่า 18.5 ภาพที่สองเป็นโรงเรียนที่มีนักเรียนน้ำหนักปกติ ค่า

BMI 18.5 – 22.9 ภาพที่สามเป็นโรงเรียนที่มีนักเรียนน้ำหนักเกิน ค่า BMI 23.0 – 24.9 ภาพที่สี่เป็นโรงเรียนที่มีนักเรียนน้ำหนักเกินขั้นอ้วน ค่า BMI 25.0 – 29.9 และภาพที่สุดท้ายเป็นโรงเรียนที่มีนักเรียนน้ำหนักเกินขั้นเป็นโรคอ้วน ค่า BMI มากกว่า 30.0

จากภาพที่ 4 เป็นการทดลองด้านรายได้ผู้ปกครองรายปี มีการกำหนดรายได้เป็น 5 ช่วง โดยภาพแรกจากซ้ายมือเป็นโรงเรียนที่มีรายได้ผู้ปกครองน้อยกว่า 40,000 บาทต่อปี ภาพที่สองเป็นโรงเรียนที่มีรายได้ผู้ปกครองอยู่ในช่วง 40,001 – 100,000 บาทต่อปี ภาพที่สามเป็นโรงเรียนที่มีรายได้ผู้ปกครองอยู่ในช่วง 100,001 – 150,000 บาทต่อปี ภาพที่สี่เป็นโรงเรียนที่มีรายได้ผู้ปกครองอยู่ในช่วง 150,001 – 300,000 บาทต่อปี และภาพที่ห้าเป็นโรงเรียนที่มีรายได้ผู้ปกครองอยู่ในช่วงมากกว่า 300,000 บาทต่อปี



ภาพที่ 4: ผลการทดลองด้านรายได้ผู้ปกครองของนักเรียน

5. สรุปผลการทดลอง

งานวิจัยนี้เป็นการสร้างภาพเสมือนสารสนเทศของนักเรียนโรงเรียนในสังกัดองค์การบริหารส่วนจังหวัดนครราชสีมา จำนวน 57 แห่ง ใช้วิธีการแบ่งกลุ่มแบบแผนผังการจัดระเบียบตัวเอง (SOM) ที่มีการนำข้อมูลสารสนเทศทางการศึกษาท้องถิ่น ได้แก่ ด้านจำนวนนักเรียน ด้านข้อมูลการเจริญเติบโตของนักเรียน และด้านรายได้ผู้ปกครองนักเรียนที่สามารถจัดสรรงบประมาณเงินอุดหนุนเฉพาะกิจให้แก่โรงเรียนในสังกัดนำไปใช้พัฒนาสถานศึกษาและช่วยเหลือนักเรียนทำให้ได้ข้อมูลแผนภาพที่สามารถนำไปช่วยวิเคราะห์จัดสรรรายการครุภัณฑ์ที่เก็บเลือดของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1 มีการนำแผนภาพด้านรายได้ผู้ปกครองมาจัดสรรเป็นรายการค่าปัจจัยพื้นฐานให้แก่นักเรียนยากจน และมีการนำข้อมูลแผนภาพด้านจำนวนนักเรียนระดับมัธยมศึกษาตอนต้นและนักเรียนระดับมัธยมศึกษาตอนปลายมาช่วยวิเคราะห์การจัดสรรงบประมาณรายการโครงการสนับสนุนการจัดการศึกษาโดยไม่เสียค่าใช้จ่าย 15 ปี เช่น ค่าเครื่องแบบนักเรียน ค่าหนังสือหนังสือแบบเรียน ค่าอุปกรณ์การเรียน ค่ากิจกรรมพัฒนาผู้เรียน เป็นต้น

เอกสารอ้างอิง

- [1] นิรันดร์ สารรัมย์. “การจัดกลุ่มโรงเรียนคุณภาพตามมาตรฐานการศึกษาโดยการใช้แผนผังการจัดระเบียบตัวเอง : กรณีศึกษาโรงเรียนสังกัดองค์การบริหารส่วนจังหวัดนครราชสีมา”การประชุมทางวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 9; 9 – 10 พฤษภาคม 2556; กรุงเทพมหานคร. กรุงเทพมหานคร : มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ; 2556. หน้า 7 – 13.
- [2] อาทิตย์ กสิบรัง. การสร้างภาพเสมือนข้อมูลพหุปัญหาเพื่อวิเคราะห์ความแตกต่างระหว่างผู้เรียน [วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์]. ขอนแก่น: บัณฑิตวิทยาลัย มหาวิทยาลัยขอนแก่น; 2550.
- [3] B. Martín-del-Brío, A. Bono-Nuez, N. Medrano-Marqués. “Self-organizing maps for embedded processor selection. Microprocessors and Microsystems” [serial online] 2005; 29(7): 307–315.

- [4] Chien.-Sing Lee, Yashwant Prasad Singh. “Student modeling using principal component analysis of SOM clusters.”*IEEE International Conference on Advanced Learning Technologies*; 30 August - 1 September 2004. [n.p.]; 2004. p. 480–484.
- [5] Melody Y. Kiang, Dorothy M. Fisher. “Selecting the right MBA schools – An application of self-organizing map networks.”*Expert Systems with Applications* [serial online] 2008; 35(3): 946–955.
- [6] TeuvoKohonen. “The self-organizing map”. *Proceedings of the IEEE* [serial online] 1990; 78(9): 1464-1480.
- [7] Ying Li, Feng Lin. “Customer segmentation analysis based on SOM clustering.”*IEEE International Conference on Service Operations and Logistics, and Informatics*; 12-15 October 2008; Beijing, China. China: IEEE; 2008. p. 15–19.

เทคนิคการสรรหาบุคลากร และคัดเลือกทีมงาน ด้วยการเลือกใช้คุณลักษณะที่เหมาะสมภายใต้พื้นฐานของสมรรถนะ A Framework of Team Recruitments with Feature Selection and Competency-based Management

จุฬาลักษณ์ วัฒนานนท์ (Julaluk Watthananon)

คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
watthananon@hotmail.com

บทคัดย่อ

ปัจจุบันการคัดเลือกบุคคล และทีมงานที่มีประสิทธิภาพเพื่อจัดทำโครงการให้บรรลุตามเป้าหมายขององค์กรถือเป็นเรื่องสำคัญ ดังนั้น ในงานวิจัยนี้ จึงได้นำเสนอเทคนิคการสรรหาบุคลากร และคัดเลือกทีมงานด้วยการเลือกใช้คุณลักษณะที่เหมาะสม ภายใต้พื้นฐานของสมรรถนะ โดยทำการทดสอบความถูกต้องในการจำแนกข้อมูลส่วนบุคคล และสมรรถนะของทีม เข้ากับสมรรถนะต้นแบบที่องค์กรคาดหวัง ผลการวิจัยพบว่า เทคนิคที่นำเสนอสามารถจำแนกบุคลากร และคัดเลือกทีมงานได้ถูกต้องตรงตามสมรรถนะต้นแบบถึง 85.02% และการคัดเลือกคุณลักษณะแบบไคสแควร์สามารถลดเวลาที่ใช้ในการสรรหาบุคลากรทดแทนภายในทีมได้ดีกว่าวิธีการอื่น

คำสำคัญ: สมรรถนะ วิธีการคัดเลือกคุณลักษณะที่เหมาะสม การคัดเลือกทีมงาน การจำแนกข้อมูล พื้นฐานของสมรรถนะ

Abstract

Currently, the team selected to achieve the goals of the organization is important. In this paper, we propose A Framework of Team Recruitments with Feature Selection and Competency-based Management. For classification accuracy test on individual competency, team competency and prototype competency by the expectations of the organization. The results show that, the technique can be classified Team Recruitments accuracy with prototype competency at 85.02% and Chi-square method can reduce the time spent on personnel in team recruitment better than the other methods.

Keyword: Competency, Feature Selection, Team Recruitment, Classification, Competency-based

1. บทนำ

ปัจจุบันจะเห็นได้ว่าหลากหลายองค์กร ต่างมีการขยายงานปรับเปลี่ยนโครงสร้างทีมงาน และรูปแบบของธุรกิจอยู่ตลอดเวลา ซึ่งส่งผลกระทบต่อสนองต่อธุรกิจ และการขยายขีดความสามารถทางการแข่งขันเพื่อสร้างความได้เปรียบ ที่จำเป็นต้องศึกษาอย่างต่อเนื่อง อันนำมาสู่การหาตัวแบบการจัดการองค์กรที่เหมาะสม [1] ซึ่งการคัดเลือกทีมงานที่เหมาะสมถือว่าเป็นปัจจัยสำคัญอย่างหนึ่ง ต่อความสำเร็จของแต่ละโครงการอย่างยิ่ง [2] โดยส่วนใหญ่วิธีการคัดเลือกทีมงานเข้าร่วมโครงการ มักระบุถึงคุณสมบัติพื้นฐานที่ทีมงานจำเป็นต้องมี อาทิ ทักษะ ความรู้ที่จำเป็น ประสิทธิภาพในการทำงาน ความมุ่งมั่น และการทำงานร่วมกับผู้อื่น เป็นต้น ซึ่งสอดคล้องกับทักษะและความสามารถในการปฏิบัติของบุคลากรในองค์กร [3], [4] จากสาระสำคัญข้างต้นหากองค์กรมีจำนวนบุคลากรที่มีคุณสมบัติดังกล่าวเป็นจำนวนมาก ซึ่งต้องนำผลการประเมินมาใช้เป็นเกณฑ์การพิจารณาคัดเลือกทีมงาน อาจก่อให้เกิดความยุ่งยากต่อการนำมาเปรียบเทียบ คัดเลือก และจัดกลุ่มของทีมงานให้มีความใกล้เคียง และเหมาะสมกับความต้องการของแต่ละโครงการ โดยจำเป็นต้องอาศัยทักษะเฉพาะของผู้เชี่ยวชาญจำแนกความแตกต่างของสมรรถนะแต่ละบุคคล เพื่อนำมาสร้างทีมงาน (Team Building) ที่มีความเหมาะสมสำหรับแต่ละโครงการให้บรรลุตามเป้าหมายที่องค์กรคาดหวัง

หลากหลายงานวิจัยมุ่งเน้นการศึกษา และพัฒนา Competence Model เพื่อใช้คัดเลือกสมรรถนะที่เหมาะสมในการพัฒนาศักยภาพของบุคคล ดังเช่นในงานวิจัยของ F. Nirschl และคณะ [5] พบว่า

ประสบการณ์ของแต่ละบุคคลเป็นพื้นฐานความสามารถในการสร้างทีมงานที่แตกต่างกัน ในขณะที่ทีมงานที่เหมาะสมในการปฏิบัติงาน ขึ้นอยู่กับขนาดของทีมงาน และการกำหนดระดับความสามารถของทีมงานที่ต้องการ ซึ่งในงานวิจัยของ C. Sang Long [7] ได้ทดสอบหาความสัมพันธ์ของปัจจัยที่ส่งผลต่อความสำเร็จอันจะนำไปสู่นักทรัพยากรมนุษย์มีอาชีพ ด้วยการเลือกใช้ 17 ปัจจัยความสามารถ งานวิจัยของ C. Hao และ F. William [9] ได้ศึกษาถึงผลกระทบของสมรรถนะทรัพยากรมนุษย์เกี่ยวกับการจัดการโครงการ ที่มีประสิทธิภาพด้วยวิธีการ SEM พบว่าผลกระทบการกำหนดสมรรถนะมีนัยสำคัญในเชิงบวก และมีอิทธิพลต่อโครงการที่มีประสิทธิภาพ ซึ่งในงานวิจัยของ A. Mahmud และคณะ [10] ได้ทำการศึกษาวิจัยในหัวข้อ สมรรถนะอะไรบ้าง คือสิ่งที่ผู้จัดการโครงการต้องการ โดยเปรียบเทียบกลุ่มตัวอย่างเข้ากับ 3 Models คือ CIOB Model, PMI Model และ CIDB Model จนได้ออกมาเป็น 4 องค์ประกอบหลักของสมรรถนะที่ควรจะมี พบว่า ผู้จัดการโครงการจำเป็นต้องมีความเชื่อมั่นในตนเอง มีพื้นฐานความรู้ที่แข็งแกร่ง และมีความสามารถในการจัดลำดับความคิด และความเข้าใจที่ดี จะสามารถควบคุมโครงการได้สำเร็จ ส่วนงานวิจัยของ K. G. Grunert และ L. Hildebrandt [11] ได้ศึกษาเกี่ยวกับปัจจัยที่นำไปสู่ความสำเร็จของการพัฒนาสมรรถนะ ผลการศึกษาอธิบายได้ว่า การปรับตัว การควบคุม และความยืดหยุ่น เป็นปัจจัยที่มีผลต่อการตอบสนองของสมรรถนะ ขณะที่ประสิทธิภาพของยุทธศาสตร์ และการสร้างแผน ที่การทำงาน เป็นปัจจัยที่สำคัญของความสำเร็จ งานวิจัยของ K. E. Dooley และ J. R. Lindner [12] ได้ทำการศึกษาและพัฒนา Model การประเมินตนเอง โดยการเลือกใช้แกนสมรรถนะของผู้เรียนเป็นหลัก พบว่า แกนสมรรถนะสามารถใช้เป็นตัวชี้วัดความคลาดเคลื่อนระหว่างกลุ่มของผู้เรียนได้ ซึ่งผลที่ได้จากการประเมินสามารถนำไปใช้ปรับกลยุทธ์การสอน และเนื้อหาในการเพิ่มความเข้มงวดของหลักสูตรได้

อีกทั้งเมื่อเร็ว ๆ นี้ในงานวิจัยของ P. Chintanaporn และ A. Mingkhwan [15] ได้นำเสนอรูปแบบของ Algorithm เพื่อเสนอแนวทางในการคัดเลือกทีมงานให้สอดคล้องกับสถานการณ์จำนวน 3 รูปแบบ ผลการศึกษาพบว่า กระบวนการ TMRC สามารถจัดทีมงานที่มีลักษณะใกล้เคียงกันได้อย่างแม่นยำ อย่างไรก็ตามยังพบปัญหาในด้านของมิติข้อมูลของสมรรถนะบางครั้ง หาก

มิติที่ใช้มีจำนวนมาก หรือมีคุณลักษณะที่ไม่เหมาะสมจะทำให้โมเดลที่ได้มีความแม่นยำน้อยลง รวมทั้งการกำหนดค่าพารามิเตอร์ที่เหมาะสมจะเป็นอีกหนทางหนึ่งในการเพิ่มประสิทธิภาพของการจัดการโครงการ [9] ซึ่งในงานวิจัยของ Weimin Chen และคณะ [6] ได้นำวิธี MARS และ CART มาใช้สำหรับการเลือก Feature ที่เหมาะสมร่วมกับ SVM พบว่าประสิทธิภาพในการจำแนกข้อมูลแม่นยำขึ้นเมื่อเทียบกับการใช้ SVM เพียงอย่างเดียว ขณะที่งานวิจัยของ วงกต และคณะ [8] ได้นำเสนอการปรับปรุงการจำแนกหมวดหมู่ของเอกสารโดยวิธีสร้างแบบจำลองหัวข้อให้กับเอกสาร โดยวิธีการคัดเลือกคุณลักษณะด้วย Information Gain และ Chi Squared พบว่า การคัดเลือกคุณลักษณะด้วยวิธี Information Gain ให้ความแม่นยำดีกว่า 10.2% ดังนั้น สำหรับงานวิจัยในครั้งนี้ ผู้วิจัยได้ให้ความสำคัญด้านการแก้ปัญหาดังกล่าว โดยการนำเสนอเทคนิคการสรรหาบุคลากร และคัดเลือกทีมงานด้วยการเลือกใช้คุณลักษณะที่เหมาะสมภายใต้พื้นฐานของสมรรถนะ เพื่อให้ได้ต้นแบบที่มีประสิทธิภาพสำหรับการพิจารณาคัดเลือกทีมงานที่เหมาะสม

2. ทฤษฎีที่เกี่ยวข้อง

2.1 สัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient)

งานวิจัยนี้ใช้การวิเคราะห์สหสัมพันธ์แบบเพียร์สัน [16] เป็นการวิเคราะห์ความสัมพันธ์ระหว่างสมรรถนะส่วนบุคคล (ตัวแปรที่ 1) และสมรรถนะต้นแบบของทีมงานที่ต้องการ (ตัวแปรที่ 2) ว่ามีความสัมพันธ์กันหรือไม่ และสัมพันธ์กันอย่างไร ความสัมพันธ์อาจเป็นไปในทิศทางเดียวกัน หรือทิศทางตรงกันข้ามได้ ซึ่งตัวแปรทิศทางของความสัมพันธ์นั้นสามารถทราบได้จากค่าสัมประสิทธิ์สหสัมพันธ์ (Coefficient of Correlation: r_{xy}) ซึ่งจะมีค่าอยู่ระหว่าง -1 กับ +1 โดยคำนวณได้จากสมการที่ (1) ดังนี้

$$r_{xy} = \frac{n \sum xy - \sum x \sum y}{\sqrt{[n \sum x^2 - (\sum x)^2] [n \sum y^2 - (\sum y)^2]}} \quad (1)$$

r_{xy} คือ ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันระหว่างตัวแปร x, y

$\sum x$ คือ ผลรวมของข้อมูลที่วัดได้จากตัวแปรที่ 1

$\sum y$ คือ ผลรวมของข้อมูลที่วัดได้จากตัวแปรที่ 2

Σ_{xy} คือ ผลรวมของผลคูณระหว่างข้อมูลตัวแปรที่ 1 และ 2

Σx^2 คือ ผลรวมของกำลังสองของข้อมูลที่วัดได้จากตัวแปรตัวที่ 1

Σy^2 คือ ผลรวมของกำลังสองของข้อมูลที่วัดได้จากตัวแปรตัวที่ 2

n คือ ขนาดของกลุ่มตัวอย่าง

ในงานวิจัยนี้ ได้มีแนวคิดว่า ค่าสมรรถนะใดที่คำนวณได้ในลักษณะพอดี และใกล้เคียงกับค่าสมรรถนะต้นแบบที่ต้องการมากที่สุด สามารถเชื่อได้ว่าบุคคลนั้นเหมาะสมกับตำแหน่งงาน ที่ได้รับมอบหมายมากที่สุด โดยผู้วิจัยได้กำหนดระดับความเชื่อมั่น ในการเลือกตัวแปรที่นำมาใช้ไว้ที่ 95% ($P < 0.05$)

2.2 การคัดเลือกคุณลักษณะ (Feature Selection) โดยผู้วิจัยทำการประเมินค่าเพื่อใช้ในการแบ่งข้อมูลด้วยการคำนวณค่า Gain สำหรับแต่ละมิติข้อมูล ดังนั้น ถ้ามิติข้อมูลใดที่ค่า Gain สูงสุด จะถูกเลือกให้เป็นกลุ่มย่อยที่มีอำนาจจำแนก สามารถคำนวณได้ดังสมการที่ (2) และ (3) [17]

$$Entropy(p) = - \sum_{i=0}^{c-1} p(j|t) \log_2 p(j|t) \quad (2)$$

โดยที่ Σ_i คือ ผลรวมของความน่าจะเป็นของค่า j ที่เกิดในทีม t

$p(j|t)$ คือ ค่าความถี่ที่มีความสัมพันธ์ของกลุ่ม j กับ โหนด t

$$Gain = Entropy(p) = - \left(\sum_{i=1}^k \frac{n_i}{n} Entropy(i) \right) \quad (3)$$

โดยที่ $Entropy(p)$ คือ ค่า Entropy ของตัวสมรรถนะต้นแบบ

$\sum_{i=1}^k \frac{n_i}{n} Entropy(i)$ คือ ค่า Entropy ในแต่ละ โหนดย่อย (Teams)

2.3 เกนเรโซ (GR) โดยทำการประเมินความน่าเชื่อถือของมิติข้อมูลสมรรถนะ โดยการวัด Gain Ratio ในแต่ละสมรรถนะ โดยใช้ค่า *SplitINFO* ดังสมการที่ (4) และการคำนวณค่าการวัด *Gain Ratio* ดังสมการที่ (5)

$$SplitINFO = \sum_{i=1}^k \frac{n_i}{n} \log_2 \frac{n_i}{n} \quad (4)$$

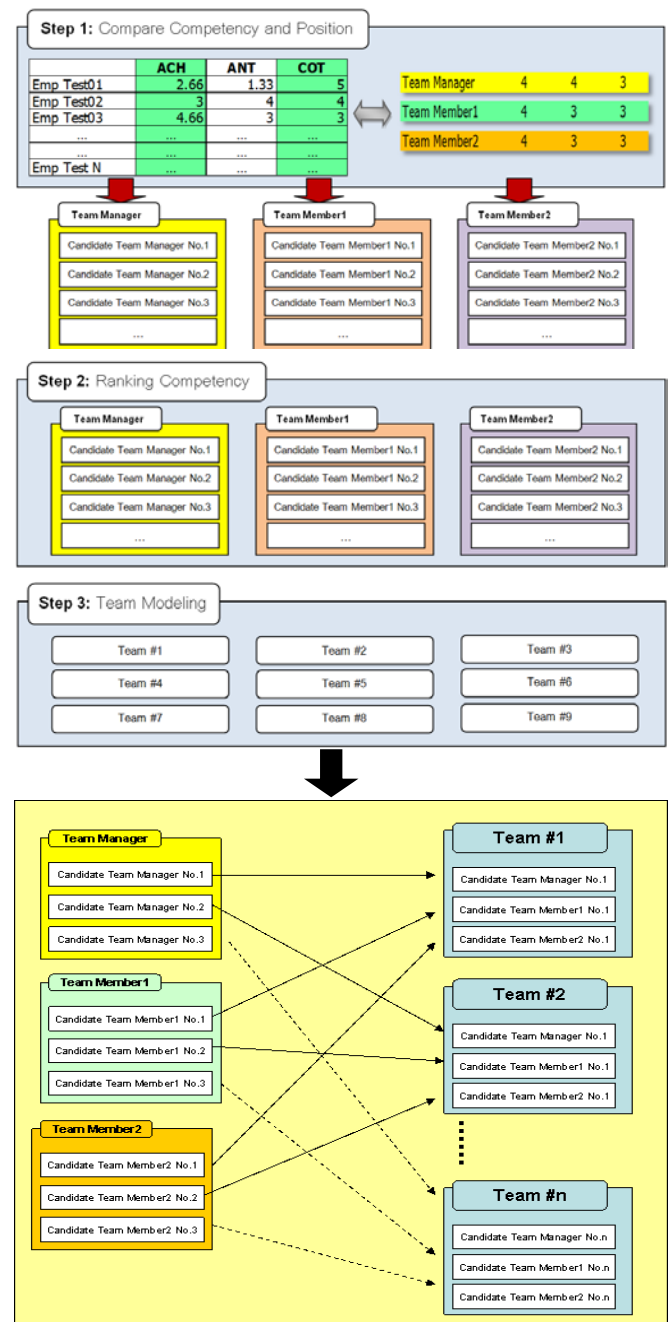
$$GainRatio = \frac{\Delta INFO}{SplitINFO} \quad (5)$$

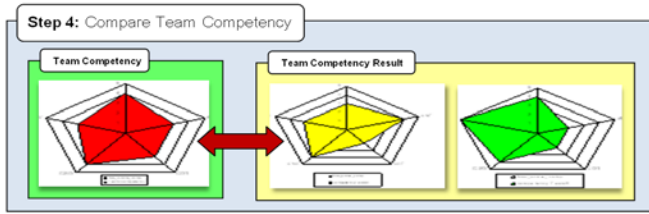
2.4 โคลสแควร์ (X^2) เป็นการประเมินค่าของ Attributes สมรรถนะ โดยใช้วิธีการคำนวณทางสถิติดังสมการที่ (6)

$$x^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} \quad (6)$$

3. วิธีการและโมเดลที่นำเสนอ

เทคนิคการสรรหาบุคลากร และคัดเลือกทีมงานด้วยการเลือกใช้คุณลักษณะที่เหมาะสม ภายใต้พื้นฐานของสมรรถนะสามารถอธิบายขั้นตอนการทำงานได้ดังภาพที่ 1





ภาพที่ 1: A Framework of Team Recruitments with Feature Selection and Competency-based Management

จากภาพที่ 1 สามารถอธิบายขั้นตอนการสรรหาและคัดเลือกทีมงานด้วยการเลือกใช้คุณลักษณะที่เหมาะสม ภายใต้พื้นฐานของสมรรถนะ ซึ่งแบ่งเป็น 4 ขั้นตอน ดังนี้

3.1 ขั้นตอนการเปรียบเทียบสมรรถนะบุคคลกับตำแหน่ง โดยแต่ละตำแหน่งจะได้รับการกำหนดจำนวนสมรรถนะที่จำเป็นอย่างเหมาะสมกับความจำเป็นของงานนั้น ๆ ซึ่งจะใช้เป็นมาตรฐานของตำแหน่งงานที่คาดหวัง (Expectation of the Performance) จากองค์กร

3.2 ขั้นตอนการจัดลำดับสมรรถนะบุคคลกับต้นแบบ โดยการใช้ค่าความคล้ายคลึง (Similarity) ใกล้เคียงมากที่สุดไปสู่วุฒิบาศที่มีค่าใกล้เคียงน้อยที่สุดตามลำดับ โดยกำหนดเงื่อนไขว่า ถ้าค่าความคล้ายคลึงที่ได้คำนวณได้เข้าใกล้ 1 มากที่สุด ให้จัดอยู่ในลำดับที่ดีที่สุด โดยจำเป็นต้องจัดลำดับจนครบทุกตำแหน่ง

3.3 ขั้นตอนการจัดทีม จากการคำนวณค่าสมรรถนะบุคคลที่มีความใกล้เคียงกับสมรรถนะต้นแบบของแต่ละตำแหน่งมาใช้งาน โดยทีมงาน (Teams) จะประกอบด้วยตำแหน่งหลายตำแหน่ง ขึ้นอยู่กับความต้องการขององค์กร จากนั้นทำการสร้าง Team Model ด้วยวิธีการคัดเลือกคุณลักษณะ (Feature Selection) และทำการจำแนกข้อมูลบุคลากรลงตามตำแหน่งงานที่สังกัดอยู่ในทีม โดยมีเงื่อนไขว่ารูปแบบของทีมต้องไม่ซ้ำกัน

3.4 ขั้นตอนจำแนกค่าความคล้ายคลึงของทีม ทำการคำนวณค่าสมรรถนะจากทีมข้างต้น จากนั้นเปรียบเทียบกับสมรรถนะต้นแบบที่องค์กรกำหนด ซึ่งกำหนดให้มีค่าอยู่ระหว่าง 0 กับ 1 โดยจะมีค่าเป็นศูนย์เมื่อทุก ๆ สมรรถนะมีผลลัพธ์เพียงแบบเดียว เช่น ใช่ทั้งหมด หรือ ไม่ใช่ทั้งหมด และจะมีค่ามากขึ้นเมื่อเริ่มมีค่าสมรรถนะที่แตกต่างกันมากขึ้น

4. การทดลองและการอภิปรายผล

4.1 ข้อมูลที่ใช้ในการทดลอง

ข้อมูลที่นำมาใช้ในการทดลองครั้งนี้ เป็นข้อมูลที่ได้จากการประเมินสมรรถนะ 10 ด้าน ของบุคลากรจำนวน 100 คน จากจำนวนพนักงานทั้งหมด 780 คน ของ บริษัท เสถียรพลาสติก แอนด์ไฟเบอร์ จำกัด โดยในงานวิจัยครั้งนี้ ผู้วิจัยเลือกใช้หลักการประเมินสมรรถนะของ Spencer [14] เพียง 10 ด้าน จากทั้งหมด 18 ด้าน เนื่องจากขอบเขตของการศึกษาครั้งนี้ เพื่อต้องการคัดเลือกทีมงาน ให้มีความเหมาะสมกับงานด้านการบริการลูกค้า (Customer Service) ซึ่งเป็นทีมงานใหม่ที่ต้องมีความต้องการคัดเลือกบุคลากรที่มีประสิทธิภาพ ภายใต้เงื่อนไขที่กำหนด โดยกำหนดสมรรถนะต้นแบบของแต่ละตำแหน่งจากความคาดหวังขององค์กร ดังแสดงตารางที่ 1

ตารางที่ 1: แสดง 10 สมรรถนะหลักที่ใช้ในการประเมิน

#	Abb.	Competency Names	TM1	TM2	TM3
1	ACH	ความมุ่งมั่นต่อความสำเร็จ	4	4	4
2	ANT	การคิดวิเคราะห์	4	3	3
3	COT	การมองภาพองค์กรรวม	3	3	3
4	CSO	การใส่ใจบริการลูกค้า	3	4	4
5	DEV	การพัฒนาผู้อื่น	4	3	3
6	IMP	การตระหนักถึงผลกระทบ	4	3	3
7	INF	การสืบค้นข้อมูล	5	4	5
8	TEL	การนำทีมงาน	5	2	2
9	TWC	ความร่วมมือและการทำงานเป็นทีม	4	4	4
10	SCF	ความเชื่อมั่นในตนเอง	4	2	2

หมายเหตุ: TM หมายถึง Team Manager

4.2 การคำนวณประสิทธิภาพ

งานวิจัยนี้เลือกใช้วิธีการวัดค่าความถูกต้อง (Accuracy) เพื่อวัดประสิทธิภาพมาตรฐาน สำหรับประเมินผลการจำแนกสมรรถนะแต่ละบุคคล และสมรรถนะทีม โดยเปรียบเทียบกับพื้นฐานสมรรถนะที่คาดหวัง โดยจำนวนของระเบียบข้อมูลที่สามารถทำนายถูกต้อง และทำนายไม่ถูกต้องด้วยตาราง Confusion Matrix ดังแสดงในตารางที่ 2

ตารางที่ 2: ตาราง Confusion matrix

		ค่าพยากรณ์			
		1	2	...	M
ค่าจริง	1	C _{1,1}	C _{1,2}	...	C _{1,M}

2	$C_{2,1}$	$C_{2,2}$...	$C_{2,M}$
...
M	$C_{M,1}$	$C_{M,2}$	$C_{M,3}$	$C_{M,M}$

กำหนดให้ $C_{i,j}$ แทนค่าที่ได้จากการทำนาย โดยที่ $i = 1, \dots, M$ และ $j = 1, \dots, M$ ถ้า $i = j$ แสดงว่าทำนายถูก (Correct Prediction) และถ้า $i \neq j$ แสดงว่าทำนายไม่ถูก (Incorrect Prediction) โดยสามารถคำนวณได้จากสมการที่ (7)

$$Accuracy = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} \quad (7)$$

4.2 การอภิปรายผล

การจำแนกบุคลากรให้มีความเหมาะสมกับตำแหน่งต่าง ๆ ในทีมงานนั้น จำเป็นต้องคำนึงถึงค่าสมรรถนะระดับบุคคลเป็นเกณฑ์เพื่อป้องกันทีมสูญเสียบุคลากรลักษณะเฉพาะบุคคลที่สำคัญไป ดังนั้น ผู้วิจัยจึงจำเป็นต้องนำค่าสมรรถนะในแต่ละด้านมาทำการเปรียบเทียบสมรรถนะแต่ละบุคคล กับสมรรถนะต้นแบบที่องค์กรต้องการรายตำแหน่ง สรุปผลดังตารางที่ 3

ตารางที่ 3: ตัวอย่างการแจกแจงบุคลากรตามตำแหน่งที่เหมาะสม

Team Manager		Team Member 1		Team Member 2	
Employee	Pearson's	Employee	Pearson's	Employee	Pearson's
Emp 20	0.74043	Emp 11	0.52943	Emp 11	0.66032
Emp 9	0.52606	Emp 12	0.31813	Emp 12	0.42700
Emp 7	0.66578	Emp 13	0.53687	Emp 13	0.58236
Emp 46	0.44215	Emp 14	0.25401	Emp 14	0.27647
Emp 41	0.47745	Emp 19	0.37237	Emp 19	0.48372
Emp 30	0.53038	Emp 2	0.24687	Emp 21	0.32970
Emp 28	0.60733	Emp 35	0.32339	Emp 35	0.44214
Emp 23	0.60076	Emp 38	0.61206	Emp 38	0.62603
Emp 22	0.45175	Emp 40	0.60072	Emp 40	0.63143
Emp 10	0.62674	Emp 8	0.22200	Emp 41	0.32193

จากนั้นทำการคัดเลือกคุณลักษณะด้วยวิธีอินฟอร์เมชันเอนโทรปี และโคสแควร์ เพื่อคัดเลือกบุคลากรที่มีสมรรถนะเฉพาะและนำมาใช้ในการจัดทีม สรุปผลดังตารางที่ 4 และ 5

ตารางที่ 4: ผลการจำแนกสมรรถนะส่วนบุคคลกับการคัดเลือกคุณลักษณะ

Filter	Feature	ACH	ANT	COT	CSO	DEV	IMP	INF	TEL	TWC	SCF
Info-Gain	100	0.771	0.747	0.807	0.76	0.752	0.807	0.754	0.738	0.808	0.752
Gain Ratio	100	0.77	0.747	0.807	0.762	0.757	0.805	0.759	0.744	0.814	0.754
Chi-square	100	0.774	0.75	0.808	0.761	0.759	0.804	0.762	0.745	0.816	0.754

ตารางที่ 5: ผลการจำแนกสมรรถนะทีมกับการคัดเลือกคุณลักษณะ

Filter	Feature	TM1	TM2	TM3
Info-Gain	1000	0.754	0.686	0.77
Gain Ratio	1000	0.758	0.693	0.771
Chi-square	1000	0.763	0.687	0.772

จากตารางที่ 4 และ 5 จะเห็นได้ว่าการคัดเลือกคุณลักษณะที่ดีคือการคัดเลือกคุณลักษณะแบบโคสแควร์ ซึ่งให้ผลการคัดเลือกโดยรวมได้รวดเร็วกว่าวิธีการอื่น

ตารางที่ 6: ผลการเปรียบเทียบสมรรถนะระหว่างทีม

	Prot_Comp	Team1	Team2	...	Team7	Team8	Team9	Rank
Prot_Comp	0	0.702	0.449	...	0.594	0.479	0.433	1
Team1	0.701	0	0.297	...	0.618	0.697	0.178	2
Team4	0.613	0.442	0.749	...	0.558	0.260	0.748	3
Team7	0.594	0.618	0.384	...	0	0.394	0.318	4
Team8	0.479	0.697	0.643	...	0.394	0	0.267	5
Team6	0.479	0.673	0.277	...	0.331	0.685	0.458	6
Team5	0.450	0.460	0.691	...	0.776	0.742	0.385	7
Team2	0.449	0.297	0	...	0.384	0.643	0.737	8
Team9	0.433	0.178	0.737	...	0.318	0.267	0	9
Team3	0.430	0.376	0.377	...	0.783	0.404	0.557	10

จากตารางที่ 6 เป็นการเปรียบเทียบความคล้ายคลึงกันระหว่างทีมจากแต่ละชุดที่แบ่งได้ถึงความคล้ายคลึงกันของสมรรถนะสูง ยกตัวอย่างเช่น ค่าสมรรถนะระหว่างทีมที่ 1 และทีมที่ 8 มีค่าเท่ากับ 0.69 หมายถึง ทั้งสองทีมมีสมรรถนะที่แตกต่างกันเท่ากับ 69% หรือมีความคล้ายคลึงกันน้อย เมื่อเทียบระหว่างทีมที่ 1 กับทีมที่ 9 ที่มีสมรรถนะคล้ายคลึงกันไปในทิศทางเดียวกัน โดยเกณฑ์ที่ยอมรับได้อยู่ที่ 0.0 - 0.4 ถือว่าเป็นค่าที่ประสิทธิภาพในการอ้างอิง [18] ดังนั้น หากอยู่ในเกณฑ์ที่กำหนดดังกล่าว องค์กรสามารถเรียกใช้บุคลากรทดแทนกันได้ ในบางตำแหน่ง โดยที่กระทบต่อผลของความคล้ายคลึงกันระหว่างสมรรถนะของทีมกับสมรรถนะต้นแบบโดยรวมน้อยที่สุด

5. สรุปผลการวิจัยและข้อเสนอแนะ

การวิจัยในครั้งนี้ มีวัตถุประสงค์นำเสนอเทคนิคการสรรหาบุคลากร และคัดเลือกทีมงานด้วยการเลือกใช้คุณลักษณะที่เหมาะสม ภายใต้พื้นฐานของสมรรถนะ ด้วยวิธีการประยุกต์ใช้เทคนิคการจำแนกข้อมูล และวิเคราะห์จัดกลุ่มข้อมูลที่เหมาะสมที่สุดด้วยสมการทางสถิติ เพื่อช่วยในการเปรียบเทียบสมรรถนะผลการวิจัยพบว่า ทีมที่สมรรถนะใกล้เคียงกับสมรรถนะต้นแบบมากที่สุด คือ ทีมที่ 1 ลำดับสองคือ ทีมที่ 4 และลำดับสาม คือ ทีมที่ 7 ซึ่งในกรณีทีมที่มีสมรรถนะใกล้เคียงกับต้นแบบมากที่สุด ไม่

สามารถปฏิบัติงานได้ด้วยสาเหตุอันใดก็แล้วแต่ องค์กรก็สามารถนำเอาผลการวิเคราะห์ที่ได้มาพิจารณา เพื่อคัดเลือกทีมลำดับที่สองมาปฏิบัติหน้าที่แทนลำดับที่หนึ่ง ได้อย่างรวดเร็ว

นอกจากนี้ ปัจจุบันองค์กรหรือหน่วยงานมักต้องเจอกับปัญหาเกี่ยวกับการเปลี่ยนแปลงขึ้นภายในทีม อาทิเช่น หัวหน้าทีมลาออกหรือบุคคลใดบุคคลหนึ่งในทีมงานไม่สามารถปฏิบัติหน้าที่ระหว่างดำเนินโครงการได้ จึงจำเป็นต้องหาบุคลากรมาทดแทนโดยด่วน ดังนั้น จากผลการทดลอง สามารถอธิบายได้ว่า การประยุกต์ใช้ค่าความคล้ายคลึงของสมรรถนะระหว่างทีมงาน ทำให้องค์กรสามารถจัดสรรบุคลากรภายในมาทดแทนตำแหน่งที่ขาดหายไปภายในทีมได้อย่างเหมาะสม และมีประสิทธิภาพ แม้ว่าผลของสมรรถนะที่ได้ของทีมจะถดถอยลง หรือน้อยกว่าการเปลี่ยนทีมงานในการรับผิดชอบโครงการทั้งทีม ซึ่งจะส่งผลกระทบต่อด้านจิตวิทยาต่อพนักงานที่เหลือในทีม ตลอดจนส่งผลให้พนักงานมีขวัญและกำลังใจในภาพรวมต่อองค์กร ไปในทิศทางที่ดี อีกทั้งบุคลากรเดิมที่อยู่ในทีมก็มีความรู้และความเข้าใจโครงการที่ได้รับมอบหมาย

ในอนาคตผู้วิจัยมีแนวคิดในการเปรียบเทียบโมเดลที่มีประสิทธิภาพสำหรับการจำแนกสมรรถนะ โดยนำเอาเทคนิคซัพพอร์ตเวกเตอร์แมชชีน ร่วมกับวิธีการคัดเลือกคุณลักษณะที่เหมาะสมทำการวิเคราะห์และประมวลผล เพื่อให้เกิดความแม่นยำในการคัดเลือกทีมให้ตรงกับความต้องการขององค์กร และครอบคลุมสถานการณ์ที่หลากหลาย ให้มีความแม่นยำมากยิ่งขึ้น

6. เอกสารอ้างอิง

- [1] ทวีเกียรติ ประพทธีระกุล และยุวัฒน์ วุฒิเมธี. “ตัวแบบการจัดการองค์กรเพื่อความได้เปรียบในการแข่งขัน ในอุตสาหกรรมไฟฟ้าและอิเล็กทรอนิกส์,” *วารสารบริหารธุรกิจ มหาวิทยาลัยธรรมศาสตร์* ปีที่ 34 ฉบับที่ 129, มกราคม – มีนาคม, พ.ศ. 2554 หน้า 24 – 40.
- [2] สุเมธ งามกนก. “การสร้างทีมงาน (Team Building),” *วารสารศึกษาศาสตร์* ปีที่ 19 ฉบับที่ 1 เดือนตุลาคม 2550 – มกราคม 2552.
- [3] นิสิตาร์ก เวชยานนท์. *บทความวิชาการด้าน HR*. กรุงเทพฯ: รัตนไตร, 2548.
- [4] นิสิตาร์ก เวชยานนท์. *Competency Model กกับการประยุกต์ใช้ในองค์กรไทย*. กรุงเทพฯ: กราฟิโกซิสเต็มส์, 2550.
- [5] F. Nirshl, M. Fuchs and J. Dorn., “A Quantitative Competence Model for e-Recruiting and Team Building in Safety Critical Domains,” *14th International Conference on Concurrent Enterprising Costa da Caparica – Lisboa, Portugal*, 23 – 35 June, 2008.
- [6] W. Chen, C. Ma and L. Ma, “Mining the customer credit using hybrid support vector machine techniques,” *Expert Systems with Applications*, vol. In Press, Uncorrected Proof, 2008.
- [7] C. Sang Long., “Examining Human Resource Competencies and Their Relationship to the Success Factors of HR Profession.,” *J. Serv. Sci. & Management*, pp. 259 – 265, 2008.
- [8] วงกต ศรีอุไร, พยุง มีสัง และชุชาติ หฤไชยะศักดิ์, “การเตรียมพีเจอาร์บนพื้นฐานแบบจำลองหัวข้อสำหรับการจำแนกหมวดหมู่ของเอกสาร”, *5th National Conference on Computing and Information Technology*, 146 – 151, 2009.
- [9] C. Hao Thi and F. William Swierczek., “The Effect of human resource competencies on project performance in Vietnamese infrastructure projects,” *Science & Technology Development*, Vol 10, No.08, 2007.
- [10] A. Mahmood, Abdulkareem M. A. Hamidaddin and M. Wira Mohd Shafiei., “What competencies do project managers need?,” *International Conference on Construction Industry (ICCI)*, 12 – 24 June, 2006.
- [11] K. G. Grunert and L. Hildebrandt., “Success factors, competitive advantage and competence development,” *Journal of Business Research* 57, pp. 459 – 461, 2004.
- [12] K. E. Dooley and J. R. Lindner., “Competencies for the Distance Education Professional: A Self-Assessment Model to Document Learning.” *28th Annual National Agricultural Education Research Conference*, December 12, 2001.
- [13] D. E. Hinkle, W. Wiersma, and S. G. Jurs, “Applied statistics for the behavioral sciences (5th Ed.),” Boston: Houghton Mifflin, 2003.
- [14] L. M., Spencer, S. M., Spencer. “Competence at Work – Models for Superior Performance,” Wiley, New York, NY, 1993.
- [15] P. Chintanaporn and A. Mingkhwan., “A Study of Information Retrieval Algorithm for Multiple Relation Competency Team Recruitment,” *Proceedings PGNet 2009, the 10th Annual Postgraduate Symposium on The Convergence of Telocommunications, Networking and Broadcasting*. Liverpool John Mores University, UK. 23-24 June 2009.
- [16] ชานินทร์ ศิลป์จารุ, *การวิจัยและวิเคราะห์ข้อมูลทางสถิติด้วย SPSS*. กรุงเทพฯ: บิสซิเนสอาร์แอนด์ดี, 2551.

- [17] P.N. Tan, M. Steinbach, and K. Vipin, *Introduction to Data Mining*. United State of America : Addison Wesley, 2005.
- [18] จุฬาลักษณ์ วัฒนานนท์ และอนิราช มิ่งขวัญ. “เทคนิคการสร้างภาพขยายความสัมพันธ์ของกลุ่มความรู้แบบสหความสัมพันธ์ระบบทศนิยมคิ้วอี่”, *การประชุมวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 4 (NCCIT09)*. กรุงเทพฯ : มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2552.

การพัฒนาระบบธุรกิจอัจฉริยะเพื่อการวางแผนการผลิตในองค์กร

The Development of Business Intelligence to Production Planning In the Organization.

ปภาดา โพธิ์คำอภิษฐ์ (Prapada Pokumapichai)¹ และมณฑิธร รัตนศิริวงศ์วุฒิ (Montean Rattanasiriwongwut)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹eyetisty@gmail.com, ²r_montean@hotmail.com

บทคัดย่อ

ปัญหาพิเศษที่จัดทำขึ้นมีวัตถุประสงค์เพื่อพัฒนาระบบธุรกิจอัจฉริยะเพื่อการวางแผนการผลิตในองค์กร เพื่อสนับสนุนการตัดสินใจของผู้บริหารในการวิเคราะห์แนวโน้มทางการตลาดและมีแนวทางในการดำเนินงานที่ชัดเจน รวมถึงการปรับกลยุทธ์ทางธุรกิจในการแข่งขันได้ดี โดยใช้หลักการธุรกิจอัจฉริยะ (Business Intelligence) ประกอบด้วยการสร้างคลังข้อมูลขนาดเล็ก (Data Mart) และนำสารสนเทศที่มีอยู่มาสร้างรายงานอัจฉริยะ (Dashboard) สามารถปรับเปลี่ยนมุมมองในการวิเคราะห์และตรงตามความต้องการของผู้บริหารและผู้ใช้งาน ผลการประเมินคุณภาพของระบบ โดยมีกลุ่มตัวอย่าง 2 กลุ่ม คือผู้เชี่ยวชาญจำนวน 5 คนและผู้ใช้งานทั่วไป จำนวน 15 คน จากผลการประเมินคุณภาพของระบบโดยผู้เชี่ยวชาญได้ค่าเฉลี่ยเท่ากับ 3.60 และค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.16 และผู้ใช้งานทั่วไปพบว่ามีค่าเฉลี่ยเท่ากับ 3.65 และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.48 ซึ่งสรุปให้เห็นว่าการพัฒนาระบบธุรกิจอัจฉริยะเมื่อนำมาประยุกต์ใช้ในการวางแผนการผลิตในองค์กรสามารถใช้งานได้จริงและในระดับดี

คำสำคัญ: ระบบธุรกิจอัจฉริยะ การวางแผนการผลิต องค์กร

Abstract

This Objective of Special Problem to develop business Intelligence to Production Planning in the Organization. to support the decision of the Management in analyzing market trends and guidelines for operation are clear. Including business strategy to compete well. Using the BI (Business Intelligence) consists of creating smaller

archives (Data Mart) and existing information and make intelligent reporting (Dashboard) can modify the view to analyze and meet the needs of Management and users. users evaluated the quality of the sample, 2 groups of 5 experts and end users including 15 of the quality evaluation system experts have an average of 3.60 and a standard deviation of 0.16 . users generally find that the mean was 3.65 and the standard deviation was 0.48 , which concluded that the development of Business Intelligence when applied in production planning in your Organization can use and that is quite good.

Keywords : Business Intelligence, Production Planning, Organization

1. บทนำ

ปัจจุบันหลายองค์กรทั้งภาครัฐบาลและเอกชนมักประสบกับปัญหาในการนำข้อมูลที่มีอยู่อย่างมหาศาลมาเปลี่ยนให้เป็นสารสนเทศที่มีคุณค่าต่อองค์กร การนำระบบสารสนเทศที่สมบูรณ์แบบมาสนับสนุนการดำเนินงานจะช่วยจัดการสายงานทุกสายงานขององค์กรให้สามารถเข้าถึงข้อมูลได้อย่างรวดเร็วและถูกต้องแม่นยำเพื่อใช้ประกอบการตัดสินใจในการดำเนินธุรกิจ

การพัฒนาระบบธุรกิจอัจฉริยะเพื่อการวางแผนการผลิตในองค์กรในครั้งนี้ มีการนำระบบจัดการฐานข้อมูลเพื่อให้ได้ข้อมูลที่นำไปสร้างความสัมพันธ์ของข้อมูลที่ต้องการ รวบรวมข้อมูลไว้รวมกัน ให้มีมาตรฐานเดียวกันและสามารถเข้าถึงข้อมูลได้โดยง่ายและแสดงออกมาในรูปแบบรายงานรายงานอัจฉริยะ (Dashboard) สามารถปรับเปลี่ยนมุมมองในการวิเคราะห์และตรงตามความต้องการของผู้บริหารและผู้ใช้งาน

เครื่องมือที่ใช้ในการวิเคราะห์ข้อมูลในหลายมิติ นั้น เป็นการประมวลผลเชิงวิเคราะห์แบบออนไลน์ (OLAP) เป็นเครื่องมือวิเคราะห์ข้อมูลและค้นหาความรู้ในคลังข้อมูลแบบหนึ่ง ที่มีความสามารถในการค้นหาและวิเคราะห์ข้อมูลจากคลังข้อมูล เพื่อตอบสนองความต้องการของผู้ใช้ได้อย่างรวดเร็ว โดยสามารถใช้ OLAP ในการวิเคราะห์ข้อมูลในลักษณะ การหมุนมิติ (Rotation) การเลือกช่วงข้อมูล (Ranging) การเลือกระดับชั้นของข้อมูล (Hierarchy)

จากข้อมูลที่ได้กล่าวมาข้างต้นมีความน่าสนใจประเด็นที่ว่าผู้บริหารจำเป็นต้องมีเครื่องมือที่ช่วยในการวิเคราะห์และการตัดสินใจ ซึ่งจะเป็นแนวทางที่จะใช้กำหนดบทบาทและภาระหน้าที่ได้อย่างชัดเจน มีประสิทธิภาพ และนำไปสู่ความสำเร็จ

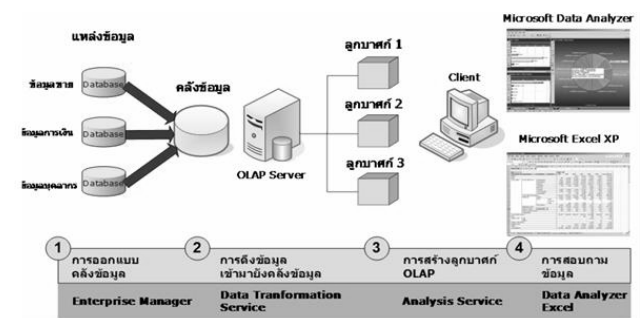
การทดสอบคุณภาพของระบบที่พัฒนาขึ้น โดยนำหลักการกระบวนการทดสอบแบบแบล็กบ็อกซ์ (Black Box Testing) เพื่อทดสอบคุณภาพของระบบ แบ่งออกเป็น 2 ขั้นตอนคือ การทดสอบในขั้นแอลฟา เป็นการทดสอบความสมบูรณ์ของระบบ โดยผู้ใช้งานทั่วไป 15 ท่าน และการทดสอบในขั้นเบต้า (Beta Testing) เป็นการทดสอบความสมบูรณ์ของระบบ โดยกลุ่มผู้เชี่ยวชาญทางด้านระบบสารสนเทศและระบบการวางแผนจำนวน 5 ท่าน การประเมินคุณภาพของระบบซึ่งได้ผลค่าเฉลี่ยเท่ากับ 3.65 และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.48 ซึ่งแสดงว่าผู้ใช้งานมีการยอมรับและมีความพึงพอใจในการใช้งานในทุกด้านอยู่ในระดับดี

2. งานวิจัยที่เกี่ยวข้อง

จากการศึกษางานวิจัยที่เกี่ยวข้องกับการพัฒนาระบบธุรกิจอัจฉริยะ สรุปได้ดังนี้

ระบบธุรกิจอัจฉริยะทางธุรกิจและการจัดการประสิทธิภาพขององค์กร ลัดดาวัลย์ [2] ได้เขียนกล่าวถึงองค์ประกอบและประโยชน์ของระบบธุรกิจอัจฉริยะ (Business Intelligence) ซึ่งประกอบด้วยสถาปัตยกรรม เครื่องมือ ฐานข้อมูล โปรแกรมการประยุกต์ใช้งาน และกรรมวิธีการใช้งานประกอบด้วย 4 องค์ประกอบ ได้แก่ แหล่งข้อมูล (data warehouse) เครื่องมือที่ใช้ดำเนินการ ค้นหา และวิเคราะห์ข้อมูล ในแหล่งข้อมูล

เครื่องมือที่ใช้ติดตาม ควบคุมและ วิเคราะห์ประสิทธิภาพ (business performance management)และท้ายสุดคือ ส่วนของผู้ใช้บริการ (user interface) การจัดการประสิทธิภาพขององค์กรและธุรกิจอัจฉริยะมีวัตถุประสงค์คือ เข้าใจแรงผลักดันของระบบธุรกิจอัจฉริยะ (BI) การเริ่มต้นในองค์กรสมัยใหม่เรียนรู้เกี่ยวกับ models ที่แตกต่างกันในการวิเคราะห์ธุรกิจ โดยยกตัวอย่างบริษัท Toyota ใช้ธุรกิจอัจฉริยะสู่ความเป็นหนึ่งเริ่มด้วยการหาปัญหา การพยายามวินิจฉัยปัญหาที่แท้จริง การศึกษาอิสระเกี่ยวกับการให้เหตุผลของการจัดการประสิทธิภาพขององค์กรและระบบ BI ก็บังคับว่า Toyota ได้บรรลุเป้าหมายแล้ว 50% ที่ได้รับผลตอบแทนจากเงินลงทุนในระบบ BI ประโยชน์ของระบบ BI ผู้จัดการต้องการข้อมูลที่ถูกต้องในเวลาและสถานที่ที่เหมาะสมแก่การทำงาน นี้คือคำกล่าวสำหรับวิธีการสมัยใหม่ของชาวปัญญาทางธุรกิจ และไม่แปลกใจที่องค์กรได้อาศัยระบบ BI เพื่อดูแลและทำความเข้าใจกิจกรรมทางธุรกิจ ดังเช่นกรณีของ Toyota ประโยชน์หลักๆของระบบ BI คือความสามารถในการจัดเตรียมสารสนเทศที่จำเป็นได้อย่างถูกต้อง รวมทั้งสามารถแสดงมุมมองรายละเอียดการปฏิบัติงานขององค์กรในรูปแบบ real-time สารสนเทศดังกล่าวจำเป็นสำหรับการตัดสินใจ การวางแผนเชิงกลยุทธ์ และการอยู่รอด [1,2]



ภาพที่ 1: สถาปัตยกรรมของคลังข้อมูล

การบริหารจัดการฐานข้อมูล วัฒนพงษ์ [3] ได้นำเสนอการบริหารจัดการฐานข้อมูล (SQL Server 2008) ประกอบด้วย 1) กระบวนการในการทำการสกัดข้อมูล ได้แก่ การรวบรวมข้อมูล แปลงรูปของข้อมูล การสร้างโครงสร้าง การประเมินโครงสร้าง การสร้างรายงาน การทำนายผล หรือให้คะแนน การรวมเข้ากับ Application และการบริหารจัดการโครงสร้าง 2) องค์ประกอบทางกายภาพของคลังข้อมูล ได้แก่ (1) Fact table

เป็นตารางศูนย์กลางที่ใช้ในคลังข้อมูล และตลาดข้อมูลที่จะเก็บจำนวน Measures และรายละเอียดสำคัญในเชิงธุรกิจ Fact Table จะเก็บ Measures ซึ่งจะเป็นตัวเลข เพราะค่าของมันเป็นพื้นฐานในการคำนวณรวมกันระหว่าง Primary Key ของ Fact Table และ Dimension Key ของ Dimension Table ได้มาจาก Multipart Keys ของ Fact table (2) Fact คือ Row ในตาราง โดย Fact จะเก็บค่าตัวเลขที่ใช้วัดเหตุการณ์ที่เกิดขึ้นกับ Data (3) Measure คือ ปริมาณ จำนวน หรือค่าทางตัวเลขซึ่งเก็บอยู่ใน Column ของ Fact Table โดย Measure จะแสดงให้เห็นถึงคุณค่าที่จะสามารถนำมาวิเคราะห์ได้ (4) Dimension คือ เอกลักษณะทางธุรกิจเป็นลักษณะทางกายภาพ (5) Dimension Table คือ ตารางในคลังข้อมูล หรือตลาดข้อมูลที่อธิบายข้อมูลใน Fact Table Dimension Tables จะประกอบด้วยรายละเอียดของข้อมูล (6) Dimension Key เป็น Foreign keys ที่เก็บไว้ใน Fact Table และ Dimension Key จะ Map กับ Primary Key ใน Dimension table จึงสามารถใช้ Surrogate Keys เป็น Primary key ของ Dimension Table 3) การสร้างรายงานจากคลังข้อมูลคือการนำข้อมูลที่ได้ออกมาเป็น Reporting Service หรือรายงานเพื่อสนองความต้องการข้อมูลในรูปแบบที่ต้องการ ซึ่งอาจจะเป็นรายงานประจำ หรือการทำรายงานแบบครั้งคราว (Ad-Hoc Report) โดยผ่านทาง Relational Database หรือ Analysis Service แม้กระทั่งข้อมูลที่เกิดจากกิจกรรมบน Web Service จะถูกแปลงเข้าไปเก็บไว้ในฐานข้อมูลแบบ Analysis Service Database โดยภายในโครงสร้างของฐานข้อมูลแบบวิเคราะห์นี้ จะทำการเก็บ Cache ของข้อมูลไว้เพื่อช่วยให้สามารถแสดงผลข้อมูลตาม Query ได้อย่างรวดเร็วหากถูกเรียกอีกในครั้งต่อไป หรือเป็น Query ชุดเดิม และมีส่วนเชื่อมต่อ Analysis Services Scripting Language (ASL) ที่สามารถติดต่อสื่อสารกับ Application อื่นด้วยภาษา XML

การวางแผนและควบคุมการผลิต วิมลสิน [6] ได้กล่าวถึงกระบวนการวางแผนและควบคุมการผลิตยุคใหม่ว่าด้วยหลักการวางแผนและควบคุมการผลิต ที่ดีคือ ความสามารถในการประสานวัตถุประสงค์ที่ขัดแย้งกันของโรงงานประกอบด้วยระดับการให้บริการลูกค้าสูงสุด(การส่งมอบสินค้าตามกำหนด) การปฏิบัติงานที่มีประสิทธิภาพด้าน โรงงาน การลงทุนในของคลังค่า วัตถุประสงค์ทั้ง 3 ประการ ดังกล่าวข้างต้นคือ

สิ่งที่ผู้บริหารฝ่ายผลิตจะต้องใช้ความพยายามในการจัดการระบบการผลิตให้บรรลุผลทุกประการ แต่โดยธรรมชาติแล้ววัตถุประสงค์ทั้ง 3 มีความขัดแย้งซึ่งกันและกัน กล่าวคือ ถ้าเราต้องการให้ต้นทุนด้านการลงทุนของคลังค่า เราจำเป็นต้องผลิตในปริมาณน้อยๆหรือเท่าที่จำเป็น แต่การดำเนินการดังกล่าวจะอาจส่งผลให้เราต้องดำเนินการผลิตสินค้ารายการเดียวกันหลายครั้ง ซึ่งในแต่ละครั้งอาจจะต้องเสียเวลาในการเตรียมการผลิต ทำให้ประสิทธิภาพของเครื่องจักรลดลง ทั้งนี้เพราะต้องสูญเสียเวลาไปกับเวลาตั้งเครื่องมากขึ้น แต่ก็อาจส่งผลให้มีสินค้าส่งมอบทันกำหนดมากขึ้น ในบางสภาพแวดล้อมของการแข่งขันทางธุรกิจ ธุรกิจจะต้องสามารถส่งมอบสินค้าในช่วงเวลาสั้นๆหรือเร็วที่สุดภายหลังจากได้รับใบสั่งจากลูกค้า กรณีดังกล่าวนี้อาจทำให้ธุรกิจจำเป็นต้องถือครองสินค้าคงคลังไว้ในปริมาณที่มากขึ้นเพื่อให้อุ่นใจในศักยภาพการตอบสนองความต้องการของลูกค้า จากที่กล่าวมาข้างต้นนี้พอจะมองเห็นได้ว่า ในการตัดสินใจด้านการวางแผนและควบคุมการผลิตจึงจำเป็นต้องคำนึงถึงผลกระทบที่จะเกิดขึ้นกับวัตถุประสงค์ตามที่กล่าวมานี้ และจะต้องพยายามทำให้ผลลัพธ์โดยรวมของวัตถุประสงค์ทั้ง 3 ประการดีที่สุด [6]

นายสันติ [4] ได้ทำการศึกษาข้อมูล BI เพื่อการวิเคราะห์ข้อมูลเพื่อการตัดสินใจด้วย Business Intelligence Data analysis for the making a decision with Business Intelligence [4]

อนุรักษ์ [5] ได้เขียนกล่าวไว้ว่าปัจจัยที่ควรคำนึงในการวางแผน แผนงบประมาณการขาย การผลิตต้องสอดคล้องกับแผนการขาย เพื่อให้แน่ใจว่ามีสินค้าพอกับการขายที่วางไว้ [5]

นฤมล [6] มีวัตถุประสงค์ในการจัดทำระบบวิเคราะห์เพื่อการบริหาร 2 ระบบคือ ระบบการวิเคราะห์ข้อมูลขาย และระบบการวิเคราะห์ความสามารถในการทำกำไรซึ่งนำเอาเทคโนโลยีการจัดเก็บข้อมูลลงในฐานข้อมูลเชิงมิติ [6]

3. วิธีการดำเนินการวิจัย

3.1 การศึกษาและรวบรวมข้อมูลระบบ จากการศึกษาระบบงานเดิม พบว่าองค์กรยังขาดการนำเทคโนโลยีสารสนเทศมาใช้ในการจัดการข้อมูล วิธีการที่จะช่วยแก้ปัญหานั้นคือการที่ต้องรวบรวมข้อมูลไว้ให้มีมาตรฐานเดียวกันและสามารถเข้าถึงข้อมูลได้โดยง่าย เพื่อสนับสนุนการตัดสินใจด้านการวางแผน

กลยุทธ์และการดำเนินการของธุรกิจ จากการศึกษาปัญหา ดังกล่าวทำให้สามารถแบ่งกระบวนการทำงานออกเป็น 2 กระบวนการใหญ่ๆ คือ กระบวนการคลังข้อมูล และกระบวนการทำรายงานทางธุรกิจเพื่อพัฒนาระบบธุรกิจอัจฉริยะเพื่อการวางแผนการผลิตในองค์กรให้มีประสิทธิภาพ และก่อให้เกิดประสิทธิผลสูงสุด

3.2 การวิเคราะห์และการออกแบบระบบ ผู้วิจัยนำข้อมูลในอดีตมาสร้างฐานข้อมูลในรูปแบบความสัมพันธ์ ผ่านโปรแกรม Microsoft SQL Server 2008 R2 และประมวลผลออกมาในรูปแบบรายงานซึ่งมีเครื่องมือให้เลือกใช้มากมาย แต่ในการศึกษาปัญหาพิเศษในครั้งนี้ผู้วิจัยเลือกใช้เครื่องมือ Business Object Version 4 เมื่อกำหนดความต้องการของระบบจากผู้ใช้งานเป็นที่เรียบร้อยแล้วในขั้นตอนต่อไปเป็นการออกแบบระบบแสดงแบบจำลองขั้นตอนการทำงาน(Data Flow) เพื่อใช้อธิบายขั้นตอนการทำงานของระบบเพื่อให้เข้าใจการทำงานของระบบมากยิ่งขึ้น

3.3 การพัฒนาระบบ การพัฒนาจะใช้โปรแกรมการจัดการฐานข้อมูล Microsoft SQL Server 2008 R2และเครื่องมือ Business Objects Version 4 ในการสร้างรายงาน ซึ่งเป็นเครื่องมือสำคัญในการพัฒนาสำหรับเรียกดูข้อมูลลูกบาศก์แบบหลายมิติ

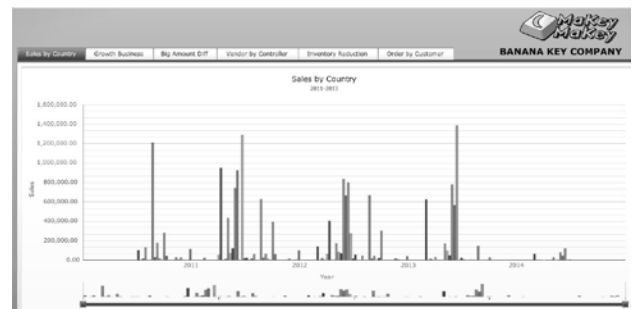
3.4 การทดสอบระบบ ผู้จัดทำได้ทำการทดสอบโดยนำหลักการกระบวนการทดสอบแบบแบล็คบ็อกซ์ (Black Box Testing) [7]โดยแบ่งการทดสอบระบบออกเป็น 2 ขั้นตอนคือการทดสอบในขั้นแอลฟา (Alpha Testing) เป็นการทดสอบความสมบูรณ์ของระบบโดยผู้ใช้งานทั่วไปจำนวน 15 ท่าน และการทดสอบในขั้นเบต้า (Beta Testing) เป็นการทดสอบความสมบูรณ์ของระบบที่จัด โดยกลุ่มผู้เชี่ยวชาญทางด้านระบบสารสนเทศและระบบการวางแผน จำนวน 5 คน

3.5 การประเมินประสิทธิภาพของระบบ เครื่องมือที่จะนำมาใช้ในการประเมินประสิทธิภาพของโครงการปัญหาพิเศษคือ แบบประเมินซึ่งมีขั้นตอนการสร้าง และกำหนดเกณฑ์ในการตัดสินใจ สถิติที่ใช้ในการวิเคราะห์ข้อมูลในครั้งนี้ ใช้สถิติเชิงพรรณนา (Descriptive Statistics) วัดค่าแนวโน้มเข้าสู่ส่วนกลาง (Measures of Central Tendency) โดยใช้ค่าเฉลี่ยเลขคณิต (Arithmetic Mean) และสถิติที่ใช้อธิบายความแตกต่างของตัว

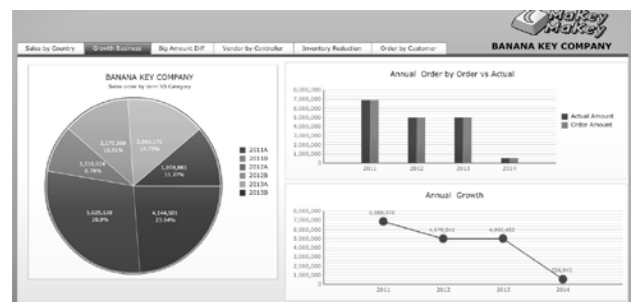
แปรที่ศึกษาโดยใช้ค่าส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) [7]

4. ผลการดำเนินงานวิจัย

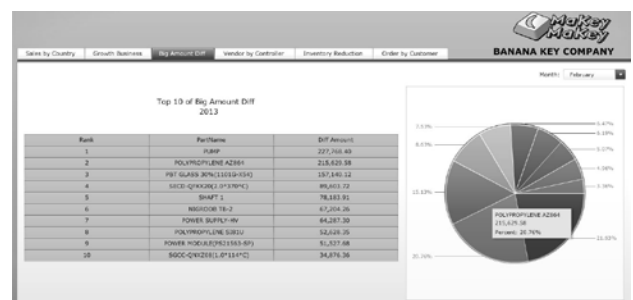
4.1 ผลการพัฒนาระบบ ผู้ใช้งานสามารถนำข้อมูลที่ได้จากคลังข้อมูลขนาดเล็ก (Data Mart) มาสร้างรายงานอัจฉริยะในรูปแบบ Dashboard สามารถปรับเปลี่ยนมุมมองในการวิเคราะห์และตรงตามความต้องการของผู้บริหารและผู้ใช้งาน



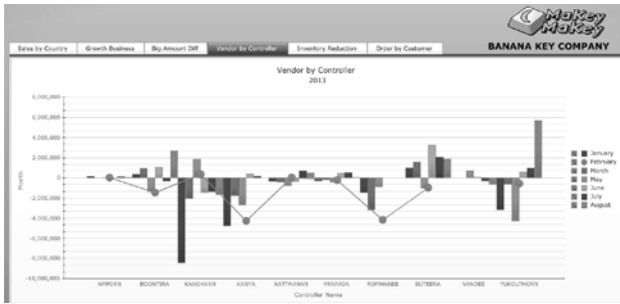
ภาพที่ 2: รายงานมูลค่าการสั่งซื้อตามภูมิภาค



ภาพที่ 3: รายงานแสดงอัตราการเติบโตของธุรกิจ



ภาพที่ 4: รายงานแสดงมูลค่าสต็อกสินค้าคงคลังคงเหลือสูงสุด 10 อันดับ



ภาพที่ 5: รายงานแสดงผลมูลค่าสต็อกสินค้าคงคลังแบ่งตามพนักงานผู้ควบคุมวัสดุปี 2013



ภาพที่ 6: รายงานวิเคราะห์สต็อกสินค้าคงคลัง 10 อันดับ



ภาพที่ 7: รายงานวิเคราะห์คำสั่งซื้อของลูกค้าแต่ละราย

จากภาพที่ 2 ถึง 7 แสดงรายงานอัจฉริยะในมุมมองที่แตกต่างกันตามแต่ความต้องการของผู้ใช้งานและผู้บริหาร อาจแบ่งเป็นมุมมองในแต่ละด้าน ได้แก่ ด้านการผลิตสินค้า สต็อกสินค้าคงคลัง และความสามารถด้านบุคลากร เป็นต้น

4.2 ผลการประเมินคุณภาพของระบบ ได้ทำการทดสอบการใช้งานจริงจากแบบประเมินและแบบสอบถามของกลุ่มผู้เชี่ยวชาญจำนวน 5 ท่าน และประชากรกลุ่มตัวอย่างจำนวน 15 ท่าน โดยแบ่งการประเมินออกเป็น 3 ด้าน คือด้านความสามารถของระบบตรงตามความต้องการ เช่น ความถูกต้องในการรายงานข้อมูลโดยรวมจากข้อมูลที่นำมาวิเคราะห์

ระบบมีความเหมาะสมของการแสดงผลข้อมูลตามที่ต้องการของแผนก ด้านความถูกต้องของระบบ เช่น ความถูกต้องในการรายงานข้อมูลโดยรวมจากข้อมูลที่นำมาวิเคราะห์และด้านการใช้งานของระบบ เช่นความง่ายในการใช้งาน

ผลการประเมินคุณภาพของระบบโดยผู้เชี่ยวชาญในทุกด้าน ดังที่กล่าวมาข้างต้น สรุปได้ว่าจากการประเมินความพึงพอใจของระบบซึ่งได้ผลค่าเฉลี่ยเท่ากับ 3.65 และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.48 ซึ่งแสดงว่าผู้ใช้งานมีการยอมรับและมีความพึงพอใจในการใช้งานในทุกด้านอยู่ในระดับดี

5. สรุปผล

งานวิจัยนี้ได้เสนอการพัฒนากระบวนการธุรกิจอัจฉริยะเพื่อการวางแผนการผลิตในองค์กรโดยใช้โปรแกรมการจัดการฐานข้อมูล Microsoft SQL Server 2008 R2 และเครื่องมือ Business Objects version 4 ในการสร้างรายงาน ระบบนี้ถือว่าเป็นส่วนหนึ่งที่ช่วยให้ผู้บริหารนำมาสนับสนุนการตัดสินใจในการวางแผนการผลิตและการวางกลยุทธ์ของธุรกิจได้อย่างมีประสิทธิภาพ ส่วนการทดสอบความถูกต้องของระบบ ใช้โปรแกรม SPSS ผู้ใช้ระบบแบ่งออกเป็น 2 กลุ่มคือ ผู้ดูแลระบบและผู้ใช้ระบบ ผลการประเมินคุณภาพของระบบโดยผู้เชี่ยวชาญในทุกด้าน พบว่ามีค่าเฉลี่ยโดยรวมทุกด้านเท่ากับ 4.02 และค่าส่วนเบี่ยงเบนมาตรฐานโดยรวมทุกด้านเท่ากับ 0.57 ผลการประเมินความพึงพอใจโดยผู้ใช้งานในทุกด้าน พบว่ามีค่าเฉลี่ยโดยรวมทุกด้านเท่ากับ 4.13 และค่าส่วนเบี่ยงเบนมาตรฐานโดยรวมทุกด้านเท่ากับ 0.62 สามารถสรุปผลโดยรวมได้ว่าระบบที่พัฒนาขึ้นมีคุณภาพและความพึงพอใจอยู่ในระดับดี ระบบสามารถนำไปใช้ในทางปฏิบัติได้อย่างมีประสิทธิภาพและสามารถตอบสนองต่อความต้องการของผู้ใช้งานอยู่ในระดับดี

เอกสารอ้างอิง

[1] เขียรชัย พัฒนศิริเวทิน, “ระบบสารสนเทศสำหรับงานการจัดการเรียนการสอนของระบบงานทวิภาคี,” Business Intelligence เบื้องต้น ปีที่ 16 ฉบับที่ 51 กรกฎาคม-กันยายน พ.ศ. 2547 หน้า 69-75.
 [2] ถัดดาวัดช์ บุญจงรักษ์, Business Intelligence และการบริหารผลการปฏิบัติงานขององค์กรและการบริหารจัดการสนับสนุนการตัดสินใจและระบบอัจฉริยะ, สาขานิติวิทยาศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น, 2553

- [3] วัฒนพงษ์ โพธิ์พั่ง, การพัฒนาระบบเพื่อวิเคราะห์ความคิดปกติของ
รายการทางธุรกิจ กรณีศึกษา: รายการทางบัญชีธนาคาร, คณะวิทยา
ศาสตรมหาบัณฑิต มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนคร
เหนือ.2550
- [4] สันติ อาริยทรัพย์และคณะ, การวิเคราะห์ข้อมูลเพื่อการตัดสินใจด้วย
Business Intelligence, คณะวิทยาการสารสนเทศ มหาวิทยาลัย
มหาสารคาม.2549
- [5] ศรีสมรภักย์ อินทจันทร์, “Business Intelligence กับการบริหาร
วางแผนและตัดสินใจ” *Productivity Forum* ปีที่ 36 ฉบับที่ 137
มกราคม-มีนาคม พ.ศ.2556 หน้า 1-5.
- [6] นฤมล พัฒนากินันท์, ระบบคลังข้อมูลเพื่อการบริหารและการ
วิเคราะห์ทางสถิติสำหรับธุรกิจนำเข้าผลิตภัณฑ์ในครัวเรือน, คณะ
พาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย. 2545
- [7] อ.ดร.วิมลสิน สุขถมยา การวางแผนและควบคุมการผลิต
(Production Planning and Control-PPC) (ie.eng.cmu.ac.th)
- [8] Thomas H. Davenport, “Competing on Analytics”, *Harvard
Business Review* harvard business review • january 2006
Page 1-12.

การพัฒนา รูปแบบการจัดการความสัมพันธ์นักศึกษาระดับบัณฑิตศึกษา

Development of a Graduate Level Student Relationship Management Model

จิรพงษ์ พงษ์พิลาสาร (Jirapong Pongpilasarn)¹ และศักดิ์ชัย ตั้งวรรณวิทย์ (Sakchai Tangwannawit)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹boykung18@hotmail.com, ²sakchait@kmutnb.ac.th

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อสังเคราะห์รูปแบบการจัดการความสัมพันธ์นักศึกษาระดับบัณฑิตศึกษา ด้วยแนวคิดจัดการความสัมพันธ์นักศึกษ โดยมิขั้นตอนนี้ 1) รวบรวมข้อมูลกิจกรรมที่เกี่ยวข้องกับการเรียนของนักศึกษา จัดทำแผนงานกิจกรรมนักศึกษา (Work Flow) จากคณะเทคโนโลยีสารสนเทศ และคณะศิลปศาสตร์ประยุกต์ 2) ทำการยืนยันกิจกรรมที่ได้ศึกษาด้วยการสัมภาษณ์เจ้าหน้าที่เชี่ยวชาญงานวิชาการทั้ง 2 คณะ ใช้การสัมภาษณ์เชิงลึก (Indepth Interview) ทำการสอบถามนักศึกษาระดับบัณฑิตศึกษา จำนวน 100 คนด้วยแบบสอบถามซึ่งได้ผ่านการหาความเชื่อมั่นด้วยวิธีของ ครอนบาค (Cronbach) 3) วิเคราะห์กิจกรรมจากข้อมูลการสัมภาษณ์และแบบสอบถามได้ฟังก์ชันกิจกรรม 4) สังเคราะห์ฟังก์ชันกิจกรรมได้อองค์ประกอบแบบจำลอง และ 5) นำองค์ประกอบแบบจำลองที่ได้นำไปประเมินความคิดเห็นจากผู้เชี่ยวชาญด้วยแบบสอบถามแบบมาตราส่วนประเมินค่า ผลการประเมินได้ค่าเฉลี่ยรวมเท่ากับ 4.08 และมีค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.58 สรุปได้ว่าแบบจำลอง SRM มีความเหมาะสมอยู่ในระดับมาก

คำสำคัญ: การจัดการความสัมพันธ์นักศึกษ กิจกรรมวิชาการ กิจกรรมการบริการ

Abstract

The objective of this study were to synthesize of a graduate level student relationship management model. The following steps were used in this research; 1) Data collection of activities student. This provided the ability to arrange a work flow of activities from samples at graduate level from the Faculty of Information Technology and

Faculty of Applied Arts. 2) Confirm using an interview of academic experts from both faculties using Indepth Interview, and 100 students from both faculties by purposive selection technique and rating-scale which is the tool for collecting the data as by academic activities questionnaire. The determined reliability of the questionnaire used the Cronbach method. 3) Analyze the activities function gained from the interviews and questionnaire. 4) Synthesize activities module obtained from the activities analysis. 5) Develop the graduate level student relationship management model. Once this was completed, the model was then evaluated by experts via questionnaires. The results of the Evaluation are $\bar{X} = 4.08$ and $S.D. = 0.58$. The Conclusion is that the SRM model is suitable in a good level.

Keywords: Student Relationship Management, Graduate Activities, Service Activities

1. บทนำ

การให้บริการด้านการศึกษาของมหาวิทยาลัยต่าง ๆ มีการแข่งขันกันสูง [1], [2] เพื่อให้บุคคลที่สนใจทั่วไปเข้ามาศึกษาต่อที่มหาวิทยาลัยของตนมากขึ้น มหาวิทยาลัยให้ความสำคัญกับการรักษาข้อมูลต่าง ๆ กิจกรรมที่มหาวิทยาลัยจัดขึ้น โดยเก็บรวบรวมข้อมูลทางการศึกษา [3], [4], [5] ที่นักศึกษาระดับบัณฑิตศึกษามีส่วนร่วมในกิจกรรมและติดต่อกับงานที่ให้บริการแก่นักศึกษา (Service) [6] เช่น การติดต่อสื่อสารระหว่างนักศึกษากับเจ้าหน้าที่ การติดตามนักศึกษาเรื่องผลการเรียน เรื่องค่าธรรมเนียมต่าง ๆ หลักสูตรแผนการเรียน ปริญญาโทและปริญญาเอก

ตารางเรียน และตารางสอบ โดยทุกกิจกรรมเป็นความสัมพันธ์ระหว่างนักศึกษา กับมหาวิทยาลัย

การสำรวจเก็บข้อมูลทำให้ทราบความต้องการของระบบที่จะช่วยอำนวยความสะดวกด้านการบริการแก่นักศึกษาระดับบัณฑิตศึกษาได้ เพิ่มความพึงพอใจของนักศึกษา

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในงานวิจัยนี้ผู้วิจัยได้ศึกษาค้นคว้าข้อมูลจากหนังสือบทความ ผู้เชี่ยวชาญ และงานวิจัยที่เกี่ยวข้องเพื่อนำไปประยุกต์ใช้ในการพัฒนารูปแบบการจัดการความสัมพันธ์นักศึกษาระดับบัณฑิตศึกษา

2.1 การจัดการความสัมพันธ์นักศึกษา (Student Relationship Management: SRM)

การจัดการความสัมพันธ์นักศึกษา หมายถึง การทำกิจกรรมต่าง ๆ การให้ความสำคัญกับการรักษาข้อมูลทางการศึกษา การสร้างความสัมพันธ์ระหว่างมหาวิทยาลัยกับนักศึกษา [1]

การจัดการความสัมพันธ์นักศึกษาศูนย์หลักการและแนวคิดจากรูปแบบการจัดการความสัมพันธ์กับลูกค้า เปลี่ยนวิธีจัดการจากลูกค้ามาเป็นนักศึกษา เปลี่ยนจากองค์กรเป็นมหาวิทยาลัย [7]

2.2 การจัดการความสัมพันธ์กับลูกค้า (Customer Relationship Management: CRM)

การจัดการความสัมพันธ์กับลูกค้า หมายถึง กลยุทธ์ทางธุรกิจเพื่อสร้างความสัมพันธ์ระยะยาว กับลูกค้า เรียนรู้ความต้องการที่แตกต่างกันของลูกค้า และตอบสนองความต้องการของลูกค้า ด้วยสินค้า หรือบริการที่เหมาะสมกับลูกค้าแต่ละคนมากที่สุด [8]

2.2.1 การจัดการความสัมพันธ์กับลูกค้าทางการศึกษา

การจัดการความสัมพันธ์กับลูกค้าทางการศึกษา ทำโดยมหาวิทยาลัยเพื่อปรับปรุงคุณภาพการศึกษาให้มีระดับสูงขึ้น โดยใช้เทคโนโลยีเข้ามาช่วยจัดการด้านการเรียนการสอน ให้มีปฏิสัมพันธ์กันมากขึ้น ช่วยอำนวยความสะดวกแก่นักศึกษา [4]

2.3 งานวิจัยที่เกี่ยวข้อง

Piedade [4] นำเสนอระบบสารสนเทศที่ใช้ในการศึกษาที่สูงขึ้น เพื่อสนับสนุนกิจกรรมทางวิชาการของนักศึกษา โดยศึกษากิจกรรมเกี่ยวกับการเรียนการสอน ศึกษาความสัมพันธ์นักศึกษากับมหาวิทยาลัย เช่น การลงทะเบียนของนักศึกษานำแนวคิดและเทคโนโลยีมาใช้ในระบบ SRM เพื่อระบุกิจกรรมของนักศึกษา เพื่อแก้ปัญหาในหมู่นักศึกษา โดยสถาบัน

การศึกษา ดูส่วนติดต่อกับนักศึกษาให้มีประสิทธิภาพเพียงพอ และติดตามผลการดำเนินงาน

ปราลี [1] นำเสนอรูปแบบระบบบริหารความสัมพันธ์นักศึกษา ศึกษาและวิเคราะห์หาปัจจัยที่มีผลต่อความพึงพอใจของนักศึกษา การทำงานของระบบ แบ่งกลุ่มข้อมูลหาความสัมพันธ์ของนักศึกษา โดยเก็บแบบสอบถามจากมหาวิทยาลัยรัฐและเอกชน แล้วจัดกลุ่มนักศึกษา หาค่าเฉลี่ยความพึงพอใจในด้านต่าง ๆ ของกลุ่มนักศึกษาแต่ละกลุ่ม

สุจิตตรา และคณะ [9] นำเสนอการติดตามสถานะของนักศึกษา ในรูปของกราฟที่สามารถปรับเปลี่ยนได้อย่างอัตโนมัติตามข้อมูลการดำเนินงานในแต่ละขั้นตอนของการทำวิทยานิพนธ์/การค้นคว้าอิสระ ข้อมูลการดำเนินงานที่ซับซ้อนในแต่ละขั้นตอน ใช้เวลามากในการตรวจสอบ จึงพัฒนาระบบขึ้นเพื่อจัดการแก้ปัญหา จากผลการทดลองใช้งานระบบพบว่าระบบที่สร้างขึ้นสามารถดำเนินงานได้เร็วขึ้น มีความถูกต้องช่วยให้ผู้ใช้ระบบสามารถติดตาม และตรวจสอบข้อมูลที่ต้องการได้ด้วยความสะดวก รวดเร็ว และข้อมูลมีความเป็นปัจจุบัน มีความถูกต้อง และเชื่อถือได้

Lechtchinskaia [2] นำเสนอการสำรวจออนไลน์วิเคราะห์ความต้องการของนักศึกษา เก็บข้อมูลจากนักศึกษาและศิษย์เก่าจาก 4 มหาวิทยาลัย มุ่งเน้นไปที่การใช้เทคโนโลยีเพื่อสนับสนุนนักศึกษาในการสื่อสารที่ทันสมัย โดยการสำรวจออนไลน์ผ่านบริการเว็บเบส (Web base) รวบรวมข้อมูลและวิเคราะห์ข้อมูลหาความพึงพอใจ งานวิจัยนี้นักศึกษามีคุณภาพความสัมพันธ์อยู่ในระดับดี

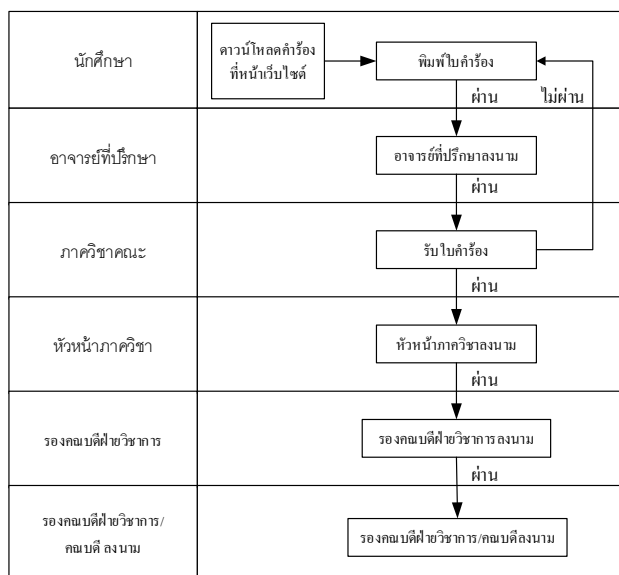
3. วิธีดำเนินการวิจัย

การพัฒนารูปแบบการจัดการความสัมพันธ์นักศึกษาระดับบัณฑิตศึกษา ดำเนินการตามขั้นตอนดังต่อไปนี้

3.1 รวบรวมข้อมูล

ศึกษางานวิชาการ ศึกษาความสัมพันธ์ระหว่างนักศึกษากับมหาวิทยาลัยจากกิจกรรมต่าง ๆ ที่คณะจัดขึ้น ศึกษางานวิจัยที่เกี่ยวข้องจากเอกสารงานวิจัย หนังสือ และเว็บไซต์ ศึกษาเครื่องมือต่าง ๆ ที่ใช้ในการพัฒนาระบบ ศึกษาเครื่องมือที่ใช้ในการสำรวจ ทฤษฎีแบบสัมภาษณ์และแบบสอบถาม [10] จัดทำแผนงานกิจกรรมนักศึกษา (Work Flow) พร้อมก็นำไป

สอบถามเจ้าหน้าที่ผู้เชี่ยวชาญงานวิชาการคณะ เพื่อตรวจสอบและแก้ไขแผนงานให้ถูกต้อง แผนงานกิจกรรมการเรียนการสอนของนักศึกษาระดับบัณฑิตศึกษามีดังต่อไปนี้ แผนงานลงทะเบียนนักศึกษาภาคปกติ ระดับปริญญาโท และปริญญาเอก แผนงานลงทะเบียนนักศึกษาภาคพิเศษ นักศึกษาเก่าตกค้างระดับปริญญาโท และปริญญาเอก แผนงานขึ้นทะเบียนนักศึกษาใหม่ภาคพิเศษ ระดับปริญญาโท และปริญญาเอก แผนงานปฐมนิเทศนักศึกษาใหม่ แผนงานการจัดทำตารางสอน และตารางสอบ แผนงานจัดทำวิทยานิพนธ์ สารนิพนธ์ และปัญหาพิเศษ แผนงานการยื่นจบการศึกษา และแผนงานการยื่นคำร้อง ตัวอย่างแผนงาน (Work Flow) กิจกรรมการเรียนดังภาพที่ 1



ภาพที่ 1: ตัวอย่างแผนงานการยื่นคำร้อง

จากภาพที่ 1 ตัวอย่างแผนงานการยื่นคำร้องมีขั้นตอนดังนี้
 1) นักศึกษาดาวน์โหลดคำร้องที่หน้าเว็บไซต์ เช่น คำร้องทั่วไป คำร้องแต่งตั้งผู้เชี่ยวชาญ คำร้องขอหนังสือขออนุญาตเก็บข้อมูลเพื่อประกอบการทำวิทยานิพนธ์/สารนิพนธ์ คำร้องขอเปลี่ยนแผนการเรียน คำร้องเพิ่ม-ถอน เป็นต้น นักศึกษาเขียนคำร้อง
 2) ส่งให้อาจารย์ที่ปรึกษาลงนาม 3) ยื่นเอกสารที่ภาควิชาคณะ ตรวจสอบคำร้อง ถ้าไม่ผ่านปรับปรุงแก้ไข และยื่นใหม่
 4) เมื่อผ่านการตรวจสอบแล้ว ภาควิชาส่งเอกสารให้หัวหน้าภาควิชา และรองคณบดีฝ่ายวิชาการ และ คณบดีลงนาม
 5) เจ้าหน้าที่ติดต่อนักศึกษารับใบคำร้องที่ภาควิชา

3.2 ยืนยันข้อมูลที่ได้ศึกษา

3.2.1 สร้างเครื่องมือด้วยแบบสัมภาษณ์ ตรวจสอบความถูกต้องของแบบสัมภาษณ์ด้วยผู้เชี่ยวชาญ โดยสัมภาษณ์เจ้าหน้าที่ผู้เชี่ยวชาญงานด้านวิชาการของคณะเทคโนโลยีสารสนเทศจำนวน 3 คน และเจ้าหน้าที่ผู้เชี่ยวชาญงานวิชาการของคณะศิลปศาสตร์ประยุกต์ จำนวน 1 คน เพื่อเก็บข้อมูลกิจกรรมนักศึกษา ใช้การสัมภาษณ์เชิงลึก (Indepth Interview) [10] โดยสัมภาษณ์จำนวน 2 ครั้ง ครั้งที่ 1 สอบถามความต้องการของระบบเกี่ยวกับกิจกรรมการเรียนของนักศึกษา ครั้งที่ 2 เพื่อทดสอบประสิทธิภาพของระบบ

3.2.2 นำข้อมูลที่ได้จากข้อ 3.2.1 สร้างแบบสอบถามเพื่อสอบถามนักศึกษาจากคณะเทคโนโลยีสารสนเทศ และคณะศิลปศาสตร์ประยุกต์ จำนวน 100 คน โดยการหาความเชื่อมั่นของแบบสอบถามทั้งฉบับด้วยวิธีของครอนบาค (Cronbach) [10] จากนักศึกษาจำนวน 30 คน คำนวณด้วยโปรแกรม SPSS ได้ผลลัพธ์ 0.86 เกินค่ามาตรฐานที่ 0.8 ขึ้นไป แสดงว่าแบบสอบถามมีความเที่ยงตรง สามารถนำไปใช้จัดเก็บข้อมูลกับนักศึกษาได้

3.3 วิเคราะห์ข้อมูล

นำข้อมูลที่ได้จากการสัมภาษณ์เจ้าหน้าที่ผู้เชี่ยวชาญงานวิชาการคณะ และข้อมูลที่ได้จากแบบสอบถามนักศึกษาถึงความต้องการของระบบสารสนเทศที่เกี่ยวกับกิจกรรมการเรียนของคณะที่ต้องการให้มีในคณะ วิเคราะห์และสรุปข้อมูล

3.4 สังเคราะห์ข้อมูล

ในขั้นตอนนี้ผู้วิจัยได้ศึกษาหาข้อมูลที่เหมาะสมกับรูปแบบการจัดการความสัมพันธ์นักศึกษาระดับบัณฑิตศึกษา (SRM) ศึกษาข้อมูลซึ่งประกอบด้วย ทฤษฎี หลักการ แนวคิด จากการสำรวจและวิเคราะห์ข้อมูลจากการสอบถาม เพื่อกำหนดกรอบแนวคิด และสังเคราะห์รูปแบบ โดยการกำหนดสภาพ และลักษณะองค์ประกอบแบบจำลองให้มีความสัมพันธ์กันอย่างถูกต้อง

3.5 ประเมินความคิดเห็นของผู้เชี่ยวชาญ

3.5.1 สร้างแบบสอบถามงานวิจัยเกี่ยวกับระดับความคิดเห็นของผู้เชี่ยวชาญที่มีต่อแบบจำลอง SRM เพื่อหาค่าความถูกต้องของแบบจำลอง ด้วยผู้เชี่ยวชาญจำนวน 5 ท่าน

3.5.2 เก็บรวบรวมความคิดเห็น นำมาวิเคราะห์ตามหลักสถิติเพื่อหาค่าดัชนีความสอดคล้อง จากนั้นปรับปรุงตามคำแนะนำ

4. ผลการดำเนินงาน

4.1 ผลการวิเคราะห์ข้อมูล

จากการสัมภาษณ์เจ้าหน้าที่ทำให้ทราบถึงกิจกรรมที่เกี่ยวข้องกับนักศึกษา ซึ่งมีกิจกรรมจำนวน 7 กิจกรรมได้แก่ 1) การติดต่อกับนักศึกษาด้วยอีเมล เว็บบอร์ด สังคมออนไลน์ (Social Network) และโทรศัพท์ 2) ติดต่ออาจารย์ที่ปรึกษาอาจารย์ผู้สอนผ่าน/ตารางนัดหมาย 3) การยื่นคำร้องผ่านระบบออนไลน์ 4) แจ้งเดือนสถานะรายวิชาที่ยังคงค้าง 5) หลักสูตรการเรียนปริญญาโทและปริญญาเอกตามสาขา 6) เก็บประวัติการลงทะเบียน ใบเสร็จ และ 7) แผนงานขั้นตอนกิจกรรมนักศึกษา ส่วนฟังก์ชันที่เกี่ยวข้องกับกิจกรรมของนักศึกษามี 4 ฟังก์ชันได้แก่ 1) ฟังก์ชันการจัดการการติดต่อ 2) ฟังก์ชันการจัดการจัดการการติดตาม 3) ฟังก์ชันการจัดการตารางเวลา 4) ฟังก์ชันการจัดการด้านเพิ่มข้อมูล ผลจากการวิเคราะห์ข้อมูลได้ดังตารางที่ 1 และ 2

จากตารางที่ 1 แสดงผลการวิเคราะห์ที่ได้จากการสัมภาษณ์เชิงลึกจากเจ้าหน้าที่เชี่ยวชาญงานด้านวิชาการทั้ง 2 คณะ ได้ลักษณะงานที่ทำงานเกี่ยวกับกิจกรรมนักศึกษา และตารางที่ 2 แสดงผลการวิเคราะห์แบบสอบถามความต้องการของนักศึกษาที่มีต่อระบบสารสนเทศที่เกี่ยวข้องกับกิจกรรมการเรียนการสอนในคณะ โดยกิจกรรมมีลักษณะงานที่เกี่ยวข้องกับโมดูล และฟังก์ชันงานของระบบที่นักศึกษาต้องการ โดยระบบที่นักศึกษาต้องการมากที่สุด คือ ระบบยื่นคำร้อง มีค่าเฉลี่ยเท่ากับ 41.67 และระบบแจ้งเดือนสถานะรายวิชาที่ยังคงค้าง มีค่าเฉลี่ยเท่ากับ 16.67

4.2 ผลการสังเคราะห์ข้อมูล

จากการวิเคราะห์กิจกรรมและฟังก์ชันงานในข้อ 4.1 ผู้วิจัยนำมาสังเคราะห์ข้อมูลได้แบบจำลองการจัดการความสัมพันธ์นักศึกษา (SRM Model) ที่มีโครงสร้างและลักษณะการทำงาน 6 องค์ประกอบดังภาพที่ 2

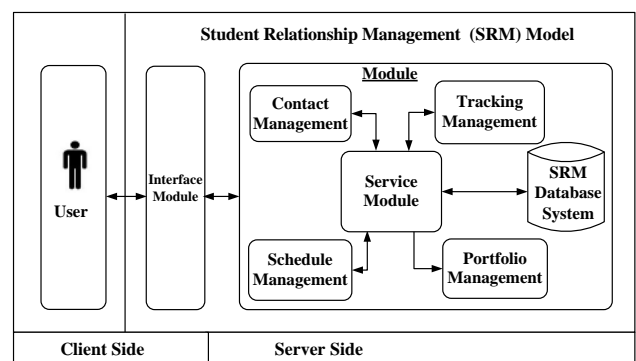
ตารางที่ 1: ผลการวิเคราะห์กิจกรรมจากแบบสัมภาษณ์เจ้าหน้าที่

ลักษณะงาน	เจ้าหน้าที่	นักศึกษา
ติดต่อกับนักศึกษาด้วยอีเมล เว็บบอร์ด สังคมออนไลน์ และโทรศัพท์	ทำช่องทางการติดต่อออนไลน์	ต้องการให้ขณะมีการติดต่อทางอีเมล
ติดต่ออาจารย์ที่ปรึกษาอาจารย์ผู้สอน/ผ่านตารางนัดหมาย	ทำตารางนัดหมาย ประสานงานระหว่างอาจารย์และนักศึกษา	ต้องการติดต่อสอบถาม อาจารย์ผ่านระบบตารางนัดหมาย
การยื่นคำร้องผ่านระบบออนไลน์	ดำเนินการจัดทำกรยื่นคำร้องผ่านระบบออนไลน์	ต้องการให้ขณะมีการยื่นคำร้องผ่านระบบออนไลน์

ลักษณะงาน	เจ้าหน้าที่	นักศึกษา
แจ้งเดือนสถานะรายวิชาที่ยังคงค้าง	จัดทำข้อมูลรายวิชา ปรับพื้นฐาน วิชาภาษาอังกฤษ ของนักศึกษา แจ้งเดือนสถานะรายวิชาที่ยังคงค้าง	ต้องการระบบแจ้งเดือนสถานะรายวิชาที่ยังคงค้าง
หลักสูตรการเรียนปริญญาโทและปริญญาเอกตามสาขา	จัดทำแผนงานหลักสูตร แยกตามภาควิชา ในคณะ	ต้องการหลักสูตรแยกตามภาควิชา ต้องการข้อมูลตารางสอน/ตารางสอบ
เก็บประวัติการลงทะเบียน และใบเสร็จ	เก็บข้อมูลการลงทะเบียนของ นศ	ต้องการประวัติการลงทะเบียน เป็นเอกสารบนเว็บไซต์
แผนงานขั้นตอนกิจกรรมนักศึกษา	จัดทำแผนงานขั้นตอนกิจกรรมนักศึกษา ทุกกิจกรรม	ต้องการขั้นตอนแผนงานกิจกรรมนักศึกษา

ตารางที่ 2: ผลการวิเคราะห์ฟังก์ชันงานจากแบบสอบถามนักศึกษา

โมดูล	ฟังก์ชันงาน	ระบบ	อ้างอิง
Contact	ส่งอีเมล รัับอีเมล และลบอีเมลได้ เว็บบอร์ด สังคมออนไลน์ และโทรศัพท์	ระบบอีเมล เว็บบอร์ด	สุวิทย์ [11], Piedade [12], Salazar [13], Catherall [14]
Contact	ติดต่ออาจารย์ทางอีเมล ตรวจสอบตารางนัดหมายอาจารย์ เก็บข้อมูลที่ได้ปรึกษาอาจารย์	ระบบตารางนัดหมาย	Lechtchinskaia [2], สุวิทย์ [11]
Tracking	ความไหลคอกสารคำร้อง ขั้นตอนการยื่นคำร้อง สถานะติดตามการยื่นคำร้อง	ระบบคำร้อง ระบบติดตาม	สุจิตรา [9]
Tracking	แสดงข้อมูลรายวิชา ปรับพื้นฐาน วิชาภาษาอังกฤษ ตรวจสอบและแจ้งเดือนสถานะ รายวิชาที่ยังคงค้าง	ระบบแจ้งเดือนสถานะรายวิชาที่ยังคงค้าง	สุจิตรา [9]
Schedule	รายละเอียดหลักสูตรปริญญาโทและปริญญาเอก	ระบบตารางนัดหมาย ตารางสอน/สอบ	ปรีชา [3], สุวิทย์ [11]
Portfolio	เก็บข้อมูลประวัตินักศึกษา เก็บประวัติการลงทะเบียน	ระบบข้อมูล	Lechtchinskaia [2], ปรีชา [3], สุวิทย์ [11]
Portfolio	แสดงแผนงานขั้นตอนกิจกรรมนักศึกษาทั้งหมด	ระบบเพิ่มข้อมูล	Lechtchinskaia [2], ปรีชา [3]



ภาพที่ 2: องค์ประกอบแบบจำลอง SRM

จากภาพที่ 2 องค์ประกอบแบบจำลอง SRM มี 6 องค์ประกอบ ดังนี้ 1) โครงสร้างและลักษณะการทำงานของส่วนการปฏิสัมพันธ์ 2) โมดูลจัดการด้านบริการแก่นักศึกษา 3) โมดูลจัดการด้านการติดต่อ 4) โมดูลจัดการด้านการติดตาม 5) โมดูลจัดการด้านตารางเวลา 6) โมดูลจัดการด้านเพิ่มข้อมูล

4.2.1 โครงสร้างและลักษณะการทำงานของส่วนการปฏิสัมพันธ์ (Interface Module)

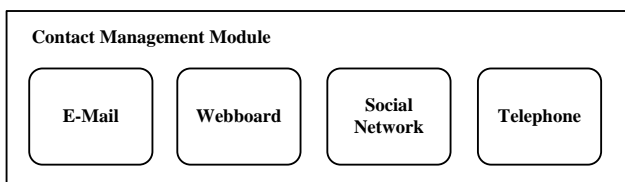
โมดูลนี้เป็นส่วนของการปฏิสัมพันธ์ การสื่อสารกับผู้ใช้ระบบ ทั้งเจ้าหน้าที่ และนักศึกษา โมดูลนี้เป็นรูปแบบของเว็บแอปพลิเคชัน (Web Application)

4.2.2 โครงสร้างและลักษณะการทำงานของโมดูลงานด้านบริการแก่นักศึกษา (Service Module)

โดยงานวิจัยนี้ใช้ระบบจัดการความสัมพันธ์ของนักศึกษา (SRM) ด้านงานบริการแก่นักศึกษา มี 4 โมดูลที่เกี่ยวข้อง ได้แก่ 1) โมดูลจัดการด้านการติดต่อ 2) โมดูลจัดการด้านการติดตาม 3) โมดูลจัดการด้านตารางเวลา 4) โมดูลจัดการด้านเพิ่มข้อมูล เพื่ออำนวยความสะดวกด้านการบริการแก่นักศึกษาระดับบัณฑิตศึกษา

4.2.3 โครงสร้างและลักษณะการทำงานของโมดูล

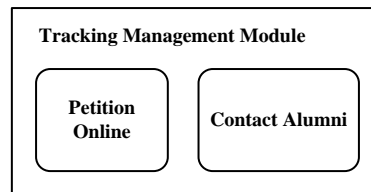
จัดการด้านการติดต่อ (Contact Management Module)



ภาพที่ 3: โครงสร้างและลักษณะการทำงานของโมดูลจัดการด้านการติดต่อ

จากภาพที่ 3 โครงสร้างและลักษณะการทำงานของโมดูลจัดการด้านการติดต่อ ใช้ติดต่อสื่อสารระหว่างนักศึกษากับเจ้าหน้าที่ และอาจารย์ ซึ่งประกอบด้วยอีเมล เว็บบอร์ด สังคมออนไลน์ (Social Network) และ โทรศัพท์

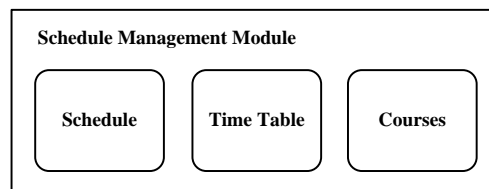
4.2.4 โครงสร้างและลักษณะการทำงานของโมดูลจัดการด้านการติดตาม (Tracking Management Module)



ภาพที่ 4: โครงสร้างและลักษณะการทำงานของโมดูลจัดการด้านการติดตาม

จากภาพที่ 4 โครงสร้างและลักษณะการทำงานของโมดูลจัดการด้านการติดตาม ใช้ติดตามนักศึกษา เรื่องการยื่นคำร้อง คำนวณโหลคำร้อง ติดตามสถานะคำร้อง ติดตามความก้าวหน้าวิทยานิพนธ์ สารนิพนธ์ และปัญหาพิเศษ ติดตามนักศึกษาศิษย์เก่า

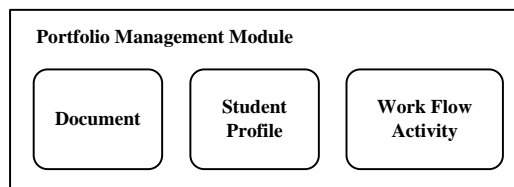
4.2.5 โครงสร้างและลักษณะการทำงานของโมดูลจัดการด้านตารางเวลา (Schedule Management Module)



ภาพที่ 5: โครงสร้างและลักษณะการทำงานของโมดูลจัดการด้านตารางเวลา

จากภาพที่ 5 โครงสร้างและลักษณะการทำงานของโมดูลจัดการด้านตารางเวลา ใช้เป็นข้อมูลตารางเวลา ตารางนัดหมาย แสดงหลักสูตรแยกตามภาควิชา แสดงวิชาปรับพื้นฐาน ภาษาอังกฤษ และแสดงตารางสอนและตารางสอบ

4.2.6 โครงสร้างและลักษณะการทำงานของโมดูลจัดการเพิ่มข้อมูล (Portfolio Management Module)



ภาพที่ 6: โครงสร้างและลักษณะการทำงานของโมดูลจัดการเพิ่มข้อมูล

จากภาพที่ 6 โครงสร้างและลักษณะการทำงานของโมดูลจัดการด้านเพิ่มข้อมูล ใช้เป็นเพิ่มข้อมูล แสดงแผนงานกิจกรรมนักศึกษา ภาควิชาศึกษาศาสตร์ และประวัติการลงทะเบียน

หลังจากนั้นนำแบบจำลองที่ได้ประเมินความคิดเห็นจากผู้เชี่ยวชาญ จำนวน 5 ท่าน ด้วยแบบสอบถามแบบมาตราส่วนประเมินค่า ซึ่งได้ผลลัพธ์คือ ค่าเฉลี่ยรวมเท่ากับ 4.08 และมีค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.58 ดังนั้นแบบจำลอง SRM มีความเหมาะสมอยู่ในระดับมาก

5. บทสรุป

งานวิจัยนี้ใช้แนวความคิดหลัก คือ แนวคิด SRM และแนวคิด CRM มาใช้ในการศึกษา แนวคิดเหล่านี้ถูกตรวจสอบผ่านด้วยการดำเนินการสัมภาษณ์และสอบถามในหลายขั้นตอนจนได้ข้อมูลที่เกี่ยวข้องกับกิจกรรมการเรียนของศึกษานำข้อมูลมาวิเคราะห์และสังเคราะห์รูปแบบและนำไปสอบถามผู้เชี่ยวชาญ ได้ค่าเฉลี่ยรวมเท่ากับ 4.08 และมีค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.58 จากนั้นปรับตามคำแนะนำจนได้รูปแบบการจัดการความสัมพันธ์นักศึกษาในระดับบัณฑิตศึกษา จากการสำรวจเก็บข้อมูลทำให้ทราบความต้องการของระบบที่จะช่วยอำนวยความสะดวกด้านการบริการแก่นักศึกษาช่วยเพิ่มความพึงพอใจของนักศึกษามากขึ้น

ในงานวิจัยต่อไปของผู้วิจัยคือ นำการสังเคราะห์แบบจำลอง SRM ที่ได้มาพัฒนาเป็นระบบการจัดการความสัมพันธ์นักศึกษาระดับบัณฑิตศึกษาต่อไป

เอกสารอ้างอิง

- [1] ปราณี มณีรัตน์, *การสร้างโมเดลจัดการระบบนักศึกษาสัมพันธ์โดยใช้เทคนิคเหมืองข้อมูล*. คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยศรีปทุม, 2011.
- [2] L. Lechtchinskaia, I. Friedrich and M.H. Breitner, "Requirements Analysis for a Student Relationship Management System – Results from an Empirical Study in Ivy League Universities," *Hawaii International Conference on System Science (HICSS)*, 2012.
- [3] ปรัชญา มั่นสันติดิน, บุญเสริม กิจศิริกุล และประสงค์ ปราณีตพลกรัง, *การบริหารความสัมพันธ์กับนักศึกษาในสถาบันระดับอุดมศึกษาโดยการประยุกต์การทำเหมืองข้อมูล*. วิทยานิพนธ์ วิทยาศาสตร์ มหาวิทยาลัยเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยศรีปทุม, 2006.
- [4] M.B. Piedade, M.Y. Santos, "Student Relationship Management: Concept, practice and technological support," *IEEE International Engineering Management Conference (IEMC)*, 2008.
- [5] B. Shannaq, Y. Rafael, and V. Alexandro, "Student Relationship in Higher Education Using Data Mining Techniques," *Global Journal of Computer Science and Technology*. Vol. 10 Issue 11, October 2010.
- [6] K. Kongsakun, C.C. Fung, S. Borirug, and W. Philuek, "An Intelligent Recommendation System Framework for Student Relationship Management," *International Conference on e-Business (iNCEB2009)*, 28 - 30 October, Bangkok, Thailand, 2009.
- [7] E. Seeman and M. O'Hara "Customer relationship management in higher education: using information systems to improve the student-school relationship," *Campus-Wide Information Systems*, Vol. 23, Issue 1, 2006.
- [8] F. Buttle, *Customer Relationship Management (Concept and Technologies)*, Taylor & Francis; 2 edition (October 14, 2008)
- [9] สุจิตรา อุดลย์เกษม, ชุติมา เจริญศิลป์, อรวรรณ เชาวลิต, จิตดำรง ปรีชาสุข, และ พิษญา พงศ์พัฒนกิจ โสคติ, "การติดตามสถานะของนักศึกษาระดับบัณฑิตศึกษาโดยใช้กราฟที่เปลี่ยนแปลงตามเวลา," *วารสารเทคโนโลยีสารสนเทศ*, 31 – 25 หน้า 14 ฉบับที่ 7 ปีที่ 1, กรกฎาคม 2, ธันวาคม –011.
- [10] มนต์ชัย เทียนทอง, *สถิติและวิธีการวิจัยทางเทคโนโลยีสารสนเทศ*. กรุงเทพมหานคร: ราชบัณฑิตยสถาน, หน้า 171-172, 2005.
- [11] สุวิทย์ มหานิยม, *ระบบสนับสนุนการประกันคุณภาพการศึกษาในส่วนของระบบการจัดการเรียนการสอนออนไลน์*. วิทยาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2009.
- [12] M.B. Piedade and M.Y. Santos, "Business intelligence in higher education: Enhancing the teaching-learning process with a SRMS," *Proceedings of the 5th Iberian Conference on Information Systems and Technologies (CISTI)*, 2010.
- [13] A. Salazar and L. Vergava, *Knowledge Discovery from E-Learning Activities*. Advance in E-Learning Experiences and Methodologies, 2008.
- [14] P. Catherall, *Delivering E-learning for Information Services in Higher Education*. Chandos Publishing. Oxford UK. 2005.

การวิเคราะห์ปัจจัยที่มีอิทธิพลต่อสมรรถนะด้านเทคโนโลยีสารสนเทศของพนักงาน ปตท.จำกัด

Affecting Analysis Factor to the IT Competence of PTT PLC.,Officers

นิธิกรณ อินทรสด (Nitikorn Intarasod)¹ และ มณเฑียร รัตน์ศิริวงศ์วุฒิ (Montean Rattanasiriwongwut)²
 สาขาวิชาระบบสารสนเทศเพื่อการจัดการ ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ
 มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
¹nitikorn.int@gmail.com , ²montean@it.kmutnb.ac.th

บทคัดย่อ

ปัญหาพิเศษนี้ เป็นการศึกษาสภาพของปัจจัยที่มีอิทธิพลต่อสมรรถนะด้านเทคโนโลยีสารสนเทศของพนักงาน ปตท.จำกัด ใช้รูปแบบการวิจัยเชิงสำรวจ เพื่อศึกษาองค์ประกอบปัจจัยที่มีอิทธิพลต่อสมรรถนะด้านเทคโนโลยีสารสนเทศของพนักงาน ปตท.จำกัด

เครื่องมือที่ใช้ในการวิจัย คือ แบบสอบถาม 2 ตอน เป็นแบบสอบถามสภาพทั่วไป และแบบสอบถามสภาพของปัจจัยที่มีอิทธิพลต่อสมรรถนะด้านเทคโนโลยีสารสนเทศของพนักงาน ปตท.จำกัด ส่วนการวิเคราะห์หองค์ประกอบเชิงสำรวจ เพื่อสำรวจจำนวนปัจจัยที่มีอิทธิพลต่อสมรรถนะด้านเทคโนโลยีสารสนเทศของพนักงาน ปตท.จำกัด ผลการวิจัยได้สมรรถนะทั้งหมด 7 ด้าน ดังนี้ ด้านความสามารถพื้นฐานเกี่ยวกับการใช้เทคโนโลยีสารสนเทศ, ด้านความรับผิดชอบต่อการใช้เทคโนโลยีสารสนเทศ, ด้านการใช้ประโยชน์ของเทคโนโลยีสารสนเทศ, ด้านความเชี่ยวชาญในการใช้เทคโนโลยีสารสนเทศ, ด้านความรู้ความเข้าใจเกี่ยวกับการใช้เทคโนโลยีสารสนเทศ, ด้านการเห็นคุณค่าในเทคโนโลยีสารสนเทศ, ด้านการประยุกต์ใช้เทคโนโลยีสารสนเทศให้เกิดประโยชน์ แล้วนำปัจจัยทั้งหมดมาทดสอบความสอดคล้องของโมเดลโดยการวิเคราะห์หองค์ประกอบเชิงยืนยัน ผลการวิจัยพบว่า โมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์เป็นอย่างดี ได้ค่า $Chi-Square = 1094.713$, $P-value = 0.194$, $GFI = 0.932$, $AGFI = 0.918$ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่าองค์ประกอบทั้ง 7 ด้าน มีอิทธิพลต่อสมรรถนะด้านเทคโนโลยีสารสนเทศของพนักงาน ปตท.จำกัด ได้เป็นอย่างดี

คำสำคัญ: สมรรถนะ การวิเคราะห์หองค์ประกอบเชิงสำรวจ, การวิเคราะห์หองค์ประกอบเชิงยืนยัน

Abstract

This Special Problems is to analyze which factor that impact to the IT competence of PTT PLC., officers by using a survey research for studying the factor composition that affect to IT competence of PTT PLC., officers.

The data collection tool is a questionnaire that divided in two parts: general information survey and impact factor of IT competence of PTT PLC., officers. SPSS was used to analyze the data in the Exploratory Factor Analysis to survey the amount of the factors which influence to the IT officers. From the research result, the overall competence were 7 cases as the using of information technology basic, the using of information technology responsibility, the using of technology for information, the using of information technology Expertise., the seeking of knowledge, the aware of information technology value, the using of Application information technology benefit. All of these cases were used to test the consistency of the model by Confirmatory Factor Analysis. The result of study found that the model consistent with Empirical data as well get the $Chi-Square = 1094.713$, $P-value = 0.194$, $GFI = 0.932$, $AGFI = 0.918$ is statistically significant at the .05 level. This value indicated that the 7 compositions affect the IT competence of PTT PLC., officers as well.

Keywords: Competence, Exploratory Factor Analysis, Confirmatory Factor Analysis

1. บทนำ

ความเจริญก้าวหน้าในยุคโลกาภิวัตน์ มีผลกระทบต่อหลายภาคส่วน อาทิ พัฒนาการด้านเทคโนโลยี และการสื่อสาร โดยเฉพาะข้อมูลสารสนเทศ เกิดการเปลี่ยนแปลงทั้งด้านเศรษฐกิจ สังคม การเมือง เทคโนโลยี และวัฒนธรรมที่เชื่อมโยงความสัมพันธ์ระหว่างบุคคล ชุมชน หน่วยธุรกิจ ภาครัฐและภาคเอกชน ทั้งในระดับประเทศ และระหว่างประเทศ ทำการเปลี่ยนแปลงอย่างรวดเร็ว ในการเข้าถึง และได้รับรู้ข้อมูลข่าวสาร หรือที่เรียกว่าโลกไร้พรมแดน ซึ่งผลจากความเจริญก้าวหน้าดังกล่าว ส่งผลให้มีการปรับเปลี่ยน และการพัฒนาเพื่อให้เกิดความเหมาะสมกับสภาพปัจจุบัน และเป็นไปตามการเปลี่ยนแปลงสอดคล้องกับนโยบายของรัฐบาล [6]

จากข้อมูลพื้นฐานดังกล่าวในขั้นตอนนี้ ผู้จัดทำจึงมุ่งศึกษาการประเมินสมรรถนะด้านเทคโนโลยีสารสนเทศที่มีผลต่อประสิทธิภาพการทำงานของพนักงานปดท.จำกัด ผลของการศึกษาในครั้งนี้ทำให้ผู้บริหาร และพนักงานเข้าใจถึงความสำคัญของการมีคุณธรรมจริยธรรม และเห็นสมควรในการอบรมสั่งสอนพนักงานให้มีความรู้ ความเข้าใจและความสามารถที่ดีต่อการใช้เทคโนโลยีสารสนเทศ ด้วยการปลูกฝังให้พนักงานเป็นคนมีจิตสำนึกถึงความรับผิดชอบ มีความรู้ความเข้าใจ เห็นคุณค่าในเทคโนโลยีสารสนเทศ และรู้จักใช้ประโยชน์ของเทคโนโลยีสารสนเทศ อันเป็นประโยชน์อย่างยิ่งต่อพนักงานในการใช้ชีวิตประจำวัน และเป็นบุคคลที่มีประสิทธิภาพต่อไป

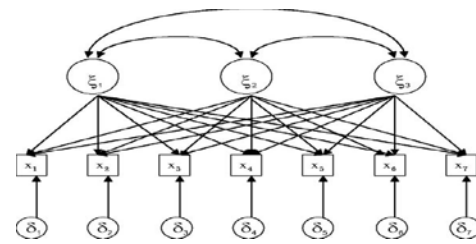
2. วรรณกรรมที่เกี่ยวข้อง

การวิเคราะห์องค์ประกอบเป็นเทคนิคที่ใช้ในการจัดกลุ่มตัวแปรประจักษ์ที่มีความเกี่ยวข้องสัมพันธ์กันให้เป็นตัวแปรแฝงหนึ่ง ๆ กลุ่มตัวแปรแฝงจะใช้ในการแทนตัวแปรประจักษ์ต่าง ๆ ตัวแปรแฝงที่ได้จากการวิเคราะห์ จะไม่มีความสัมพันธ์กัน (Multi Collinearity) ตัวแปรแฝงจะถือเสมือนเป็นหนึ่งตัวแปรใหม่ ประโยชน์ของการวิเคราะห์องค์ประกอบ ได้แก่ การใช้ในการลดความสัมพันธ์กันระหว่างตัวแปรประจักษ์ที่ทำให้เกิดความผิดพลาดในการวิเคราะห์ข้อมูลต่างๆ เช่นการวิเคราะห์สมการถดถอย และการจัดกลุ่มของตัวแปรประจักษ์ต่อตัวแปรแฝงเพื่อนำไปประมวลผลหาความสัมพันธ์ของตัวแปรแฝงต่างๆ ว่ามีความเกี่ยวข้องกันเช่นไรในการวิเคราะห์สมการ

โครงสร้าง (Structural equation modeling) ต่อไป [2], [6]

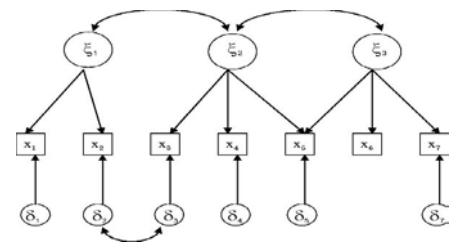
2.1 วัตถุประสงค์ของการวิเคราะห์องค์ประกอบ
วัตถุประสงค์ในการวิเคราะห์องค์ประกอบมี 2 ประการ คือ

2.1.1 เพื่อสำรวจหรือค้นหาตัวแปรแฝง ที่ซ่อนอยู่ภายใต้ตัวแปรที่สังเกตได้หรือวัดได้เรียกว่าการวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis : EFA) แสดงดังภาพที่ 1 [6], [7]



ภาพที่ 1: โมเดลการวิเคราะห์องค์ประกอบเชิงสำรวจ

2.1.2 เพื่อทดสอบความสอดคล้องหรือยืนยันทฤษฎีที่ผู้เฝ้าค้นพบเรียกว่า การวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Analysis : CFA) แสดงดังภาพที่ 2 [7]



ภาพที่ 2: โมเดลการวิเคราะห์องค์ประกอบเชิงยืนยัน

2.2 ประเภทของเทคนิคการวิเคราะห์องค์ประกอบ

เทคนิคของการวิเคราะห์องค์ประกอบ แบ่งออกเป็น 2 ประเภทคือ

2.2.1 การวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis) การวิเคราะห์องค์ประกอบเชิงสำรวจจะใช้ในกรณีที่ผู้ศึกษาไม่มีความรู้ หรือมีความรู้ไม่มากนักเกี่ยวกับ โครงสร้างความสัมพันธ์ของตัวแปรเพื่อศึกษาโครงสร้างของตัวแปร และลดจำนวนตัวแปรที่มีอยู่เดิมให้มีการรวมกันได้

2.2.2 การวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Analysis) การวิเคราะห์องค์ประกอบเชิงยืนยันจะใช้กรณีที่ผู้ศึกษาทราบ โครงสร้างความสัมพันธ์ของตัวแปร หรือคาดว่า โครงสร้างความสัมพันธ์ของตัวแปรควรจะเป็นรูปแบบใด หรือคาดว่าตัวแปรใดบ้างที่มีความสัมพันธ์กัน

มากและควรรู้อยู่ในองค์ประกอบเดียวกัน หรือคาดว่ามีความแปรปรวนที่ไม่มีความสัมพันธ์กัน ควรจะอยู่ต่างองค์ประกอบกัน หรือกล่าวได้ว่า ผู้ศึกษาทราบโครงสร้างความสัมพันธ์ของตัวแปรหรือคาดว่าโครงสร้างความสัมพันธ์ของตัวแปรเป็นอย่างไร และจะใช้เทคนิคการวิเคราะห์องค์ประกอบเชิงยืนยันมาตรวจสอบหรือยืนยันความสัมพันธ์ว่าเป็นอย่างไรหรือไม่ โดยการวิเคราะห์หาความตรงเชิงโครงสร้างนั่นเอง [6]

2.3 งานวิจัยที่เกี่ยวข้อง

ไชยา เข็นแฆ (2547) ศึกษาเรื่อง ความรู้ความเข้าใจ และทัศนคติ ที่มีผลต่อพฤติกรรมการทำงานด้วยระบบเทคโนโลยีสารสนเทศของผู้ปฏิบัติงานฝ่ายปฏิบัติการภาคกลาง การไฟฟ้าฝ่ายผลิตแห่งประเทศไทย ผลการศึกษาพบว่า ความรู้ความเข้าใจด้านเทคโนโลยีสารสนเทศของผู้ปฏิบัติงานมีความสัมพันธ์กับทัศนคติของผู้ปฏิบัติงาน ฝากต่อการทำงานด้วยระบบเทคโนโลยีสารสนเทศของผู้ปฏิบัติงาน ฝากที่มีความรู้ความเข้าใจด้านเทคโนโลยีสารสนเทศสูงจะมีทัศนคติที่เห็นด้วยต่อการความคิดเห็นเรื่อง เทคโนโลยีสารสนเทศทำให้ได้รับประโยชน์มากขึ้น อายุ ระดับการศึกษา และอายุงานที่แตกต่างกันมีผลต่อพฤติกรรมการทำงานด้วยระบบเทคโนโลยีสารสนเทศในด้านการใช้งานคอมพิวเตอร์การทำงานด้วยระบบเครือข่าย และการทำงานด้วยอินเทอร์เน็ต [1]

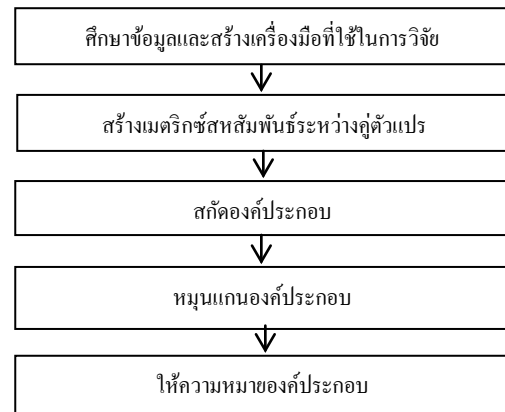
นิพนธ์ ลือภักดินันท์ (2547) ศึกษาเรื่อง ปัจจัยที่มีผลต่อความสำเร็จขององค์กรในการใช้ระบบสารสนเทศในการปฏิบัติงานของพนักงานบริษัท คาโอ อินดัสตรีล (ประเทศไทย) จำกัด ผลการศึกษาพบว่า พนักงานส่วนใหญ่มีความรู้ความเข้าใจในระบบสารสนเทศ และมีความเห็นด้วยกับลักษณะระบบสารสนเทศในด้านความน่าเชื่อถืออยู่ในระดับปานกลาง ในด้านความปลอดภัยและด้านความสะดวกอยู่ในระดับมาก พนักงานมีความเห็นด้วยอยู่ในระดับมาก กับความสำเร็จ [3]

3. วิธีการดำเนินการวิจัย

การดำเนินงานวิจัย เรื่องการวิเคราะห์ปัจจัยที่มีอิทธิพลต่อสมรรถนะด้านเทคโนโลยีสารสนเทศของพนักงาน ปตท.จำกัด โดยการวิเคราะห์องค์ประกอบเชิงสำรวจ ผู้วิจัยได้กำหนดขั้นตอนของการดำเนินงาน ออกเป็นรายละเอียดดังนี้

3.1 การวิเคราะห์องค์ประกอบเชิงสำรวจ

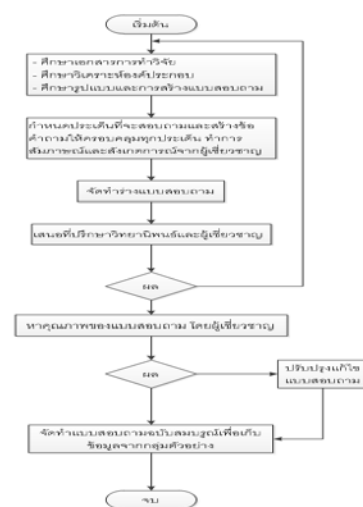
การวิเคราะห์องค์ประกอบเชิงสำรวจ เป็นการสำรวจจำนวนองค์ประกอบ ที่แฝงอยู่ภายใต้ ตัวแปรที่สังเกตได้ เพื่อกำหนดจำนวนองค์ประกอบที่ได้จากการวิเคราะห์และอธิบายความแปรปรวนที่สังเกตได้ เป็นขั้นตอนการวิเคราะห์องค์ประกอบเชิงสำรวจ สามารถแบ่งการวิเคราะห์ออกเป็น 5 ขั้นตอน แสดงรายละเอียดดังภาพที่ 3 [3]



ภาพที่ 3: ขั้นตอนการวิเคราะห์องค์ประกอบเชิงสำรวจ (EFA)

3.1.1 ศึกษาข้อมูลและสร้างเครื่องมือที่ใช้ในการวิจัย

ผู้วิจัยได้ทำการศึกษาค้นคว้างานวิจัยที่เกี่ยวข้อง เอกสารและตำราต่างๆที่เกี่ยวข้องกับสมรรถนะของนักเทคโนโลยีสารสนเทศ รูปแบบของการวิเคราะห์องค์ประกอบและรูปแบบการสร้างแบบสอบถาม จากนั้นได้กำหนดประเด็นที่ตั้งข้อคำถามและร่างข้อคำถามให้ครอบคลุมทุกประเด็น นำเสนอให้อาจารย์ที่ปรึกษาพิจารณาแล้วนำมาจัดทำแบบสอบถามฉบับร่าง การดำเนินงานได้แสดงดังภาพที่ 4 [7]



ภาพที่ 4: ขั้นตอนการสร้างและหาคุณภาพแบบสอบถาม

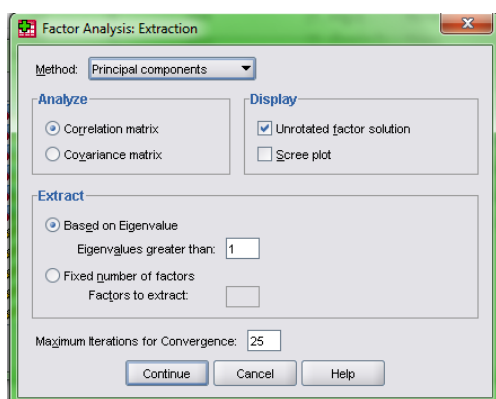
ซึ่งสามารถรวบรวมองค์ประกอบหลักได้ 7 ด้านมีข้อความทั้งหมด 40 ข้อ แล้วนำไปหาคุณภาพของแบบสอบถามโดยขอความคิดเห็นจากผู้เชี่ยวชาญจำนวน 3 ท่าน แล้วนำค่าคำนวณหา ค่า IOC ซึ่งข้อความใช้ได้ทั้งหมด จึงนำไปจัดทำแบบสอบถามฉบับสมบูรณ์และนำไปหาค่าความเชื่อมั่น ของแบบสอบถาม จากกลุ่มตัวอย่างจำนวน 30 คน ซึ่งผลการคำนวณ ข้อคำถามทุกด้านมีค่าความเชื่อมั่น สามารถนำไปใช้งานได้

3.1.2 การสร้างเมตริกซ์สหสัมพันธ์ระหว่างคู่ตัวแปร

การสร้างเมตริกซ์ความสัมพันธ์ระหว่างคู่ตัวแปรทุกตัวเป็นขั้นตอนแรกของการวิเคราะห์องค์ประกอบ ที่จะดำเนินการหาความสัมพันธ์ในรูปแบบเส้นตรงโดยวิธีของ Pearson Correlation ระหว่างตัวแปรทุกคู่ที่ต้องการนำมาจัดกลุ่มกันซึ่งอยู่ในรูปของเมตริกซ์สหสัมพันธ์ระหว่างคู่ของตัวแปร เพื่อที่จะลดขนาดของข้อมูลให้น้อยลงเพื่อที่จะสะดวกในการสกัดองค์ประกอบ

3.1.3 การสกัดองค์ประกอบ (Factor Extraction)

การสกัดองค์ประกอบ เป็นการหาจำนวนองค์ประกอบ ที่สามารถใช้ตัวแปรทั้งหมดทุกตัวได้ การสกัดองค์ประกอบมีหลายวิธีแต่ละวิธีจะให้ผลแตกต่างกัน ในการวิเคราะห์องค์ประกอบในงานวิจัยนี้เลือกใช้การสกัดองค์ประกอบแบบองค์ประกอบหลัก (Principal Component Analysis) วิธีนี้อาศัยหลักความสัมพันธ์เชิงเส้นระหว่างตัวแปรที่ใช้เป็นข้อมูลแสดงดังภาพที่ 5

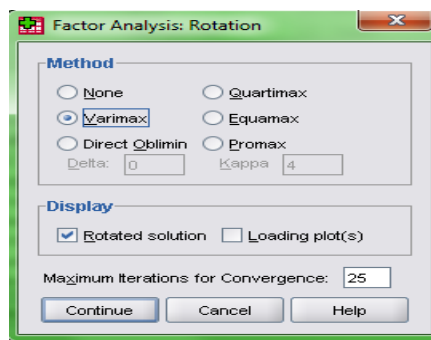


ภาพที่ 5: การกำหนดวิธีการสกัดปัจจัย

3.1.4 การหมุนแกนองค์ประกอบ (Factor Rotation)

การหมุนแกนองค์ประกอบ เป็นขั้นตอนการแยกตัวแปรให้เด่นชัดว่าตัวแปรตัวหนึ่งๆ ควรจะจัดอยู่ในองค์ประกอบใด

เนื่องจากการสกัดองค์ประกอบ จะได้องค์ประกอบหลายองค์ประกอบ ซึ่งแต่ละองค์ประกอบจะเกิดการรวมของตัวแปรเชิงเส้นตรง ในการวิเคราะห์องค์ประกอบในงานวิจัยนี้ เลือกใช้การหมุนแกนตัวประกอบแบบ Varimax จำนวน 25 รอบดังแสดงดังภาพที่ 6



ภาพที่ 6: การหมุนแกนองค์ประกอบ

3.1.5 การให้ความหมายขององค์ประกอบ (Factor Meaning)

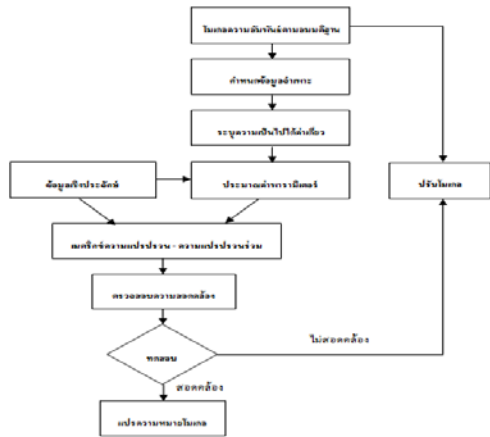
การให้ความหมายแก่องค์ประกอบเป็นขั้นตอนที่จะต้องกำหนดชื่อให้ความหมายแก่องค์ประกอบหรือตัวแปรที่มีความสัมพันธ์กัน โดยพิจารณาองค์ประกอบนั้นๆ ประกอบด้วยตัวแปรอะไรเป็นสมาชิก อยู่ในแต่ละองค์ประกอบ มีน้ำหนักความสำคัญของการเป็นสมาชิกแตกต่างกัน ซึ่งการตั้งชื่อองค์ประกอบ ควรให้สอดคล้องกับตัวแปรในองค์ประกอบนั้นๆ มากที่สุด แสดงดังตารางที่ 1

ตารางที่ 1: การให้ความหมายขององค์ประกอบ และแสดงค่าความเชื่อมั่นแต่ละด้าน

องค์ประกอบ / ตัวแปร	จำนวนข้อ	ค่าความเชื่อมั่นของแบบสอบถาม
ด้านความรู้ความเข้าใจเกี่ยวกับการใช้เทคโนโลยีสารสนเทศ	3	.866
ด้านการประยุกต์ใช้เทคโนโลยีสารสนเทศให้เกิดประโยชน์	3	.861
ด้านการเห็นคุณค่าในเทคโนโลยีสารสนเทศ	6	.857
ด้านความเชื่อวางใจในการใช้เทคโนโลยีสารสนเทศ	4	.849
ด้านความสามารถพื้นฐานเกี่ยวกับการใช้เทคโนโลยีสารสนเทศ	7	.848
ด้านการใช้ประโยชน์ของเทคโนโลยีสารสนเทศ	5	.845
ด้านความรับผิดชอบต่อการใช้เทคโนโลยีสารสนเทศ	8	.844

3.2 การวิเคราะห์องค์ประกอบเชิงยืนยัน

เป็นขั้นตอนในการตรวจสอบและยืนยันความสอดคล้องของผลการวิเคราะห์องค์ประกอบเชิงสำรวจตามทฤษฎีกับข้อมูลเชิงประจักษ์สามารถแบ่งขั้นตอนการวิเคราะห์องค์ประกอบเชิงยืนยันได้แสดงดังภาพที่ 7 [5] , [4]



ภาพที่ 7: ขั้นตอนการวิเคราะห์องค์ประกอบเชิงยืนยัน

3.2.1 วิเคราะห์สเกลองค์ประกอบ นำผลจากการวิเคราะห์องค์ประกอบเชิงสำรวจ ซึ่งประกอบด้วยค่าทางสถิติต่างๆของตัวแปรที่สังเกตได้ในโมเดลการวิจัย ประกอบด้วยค่าเฉลี่ย ค่าส่วนเบี่ยงเบนมาตรฐานและเมตริกสหสัมพันธ์ระหว่างคู่ตัวแปรมาทำการสร้างสเกลขององค์ประกอบแต่ละด้านโดยพิจารณาว่าโมเดลตามทฤษฎีที่ได้จากการวิเคราะห์องค์ประกอบเชิงสำรวจสอดคล้องกับข้อมูลเชิงประจักษ์ที่ได้จากการเก็บรวบรวมข้อมูลจากกลุ่มตัวอย่าง หากโมเดลไม่สอดคล้องกันก็สามารถปรับสเกลของโมเดลช่วยจนสอดคล้องกันได้

3.2.2 การวิเคราะห์องค์ประกอบด้วยโปรแกรมลิสเรล สร้างโมเดลวัดองค์ประกอบซึ่งเริ่มจากการกำหนดข้อมูลจำเพาะของโมเดล ระบุความเป็นไปได้ของโมเดล และทำการประมาณค่าพารามิเตอร์ในโมเดล

3.2.3 การแปลผลการวิเคราะห์องค์ประกอบ การแสดงผลการวิเคราะห์โมเดลการวัด เพื่ออธิบายการทดสอบความสอดคล้องของโมเดลการวัดขององค์ประกอบด้านต่างๆ ที่จะดูว่าโมเดลตามทฤษฎีที่ได้จากการวิเคราะห์องค์ประกอบเชิงสำรวจสอดคล้องกับข้อมูลเชิงประจักษ์ที่ได้จากการเก็บรวบรวมข้อมูลจากกลุ่มตัวอย่าง โดยพิจารณาค่า chi-square, P-value, ค่าดัชนีวัดระดับความกลมกลืน (GFI) ค่าดัชนีวัดระดับความกลมกลืนที่ปรับแก้แล้ว (AGFI) แสดงว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์เป็นอย่างดี และได้โมเดลรูปแบบความสัมพันธ์เชิงสาเหตุ โดยการวิเคราะห์องค์ประกอบเชิงยืนยันลำดับสอง และค่าน้ำหนักองค์ประกอบทั้งหมดมีค่าเป็นบวก มีนัยสำคัญทางสถิติที่ระดับ 0.05 ทุกองค์ประกอบ

โดยน้ำหนักขององค์ประกอบในแต่ละด้านเรียงตามความสำคัญขององค์ประกอบ

4. ผลการดำเนินงานวิจัย

ผลการดำเนินงานการวิเคราะห์ปัจจัยที่มีอิทธิพลต่อสมรรถนะด้านเทคโนโลยีสารสนเทศของพนักงาน ปตท.จำกัด โดยใช้วิธีการวิเคราะห์องค์ประกอบเชิงยืนยันผู้วิจัยได้แบ่งผลการวิจัยออกเป็น 2 ส่วน ดังต่อไปนี้

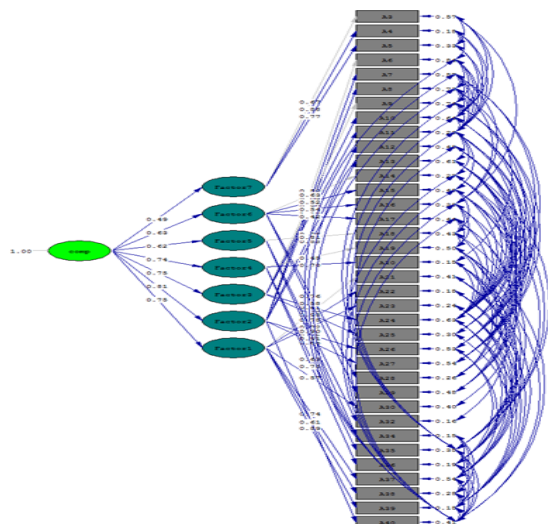
4.1 การวิเคราะห์องค์ประกอบเชิงสำรวจ

ผลการวิจัยการวิเคราะห์องค์ประกอบเชิงสำรวจ ปัจจัยแต่ละด้าน คือ ดังนี้ ด้านความรู้ความสามารถเกี่ยวกับการใช้เทคโนโลยีสารสนเทศ, ด้านความรับผิดชอบต่อการใช้เทคโนโลยีสารสนเทศ, ด้านการใช้ประโยชน์ของเทคโนโลยีสารสนเทศ, ด้านความเชี่ยวชาญในการใช้เทคโนโลยีสารสนเทศ, ด้านการแสวงหาความรู้, ด้านการเห็นคุณค่าใน

เทคโนโลยีสารสนเทศ, ด้านการประยุกต์ใช้เทคโนโลยีสารสนเทศให้เกิดประโยชน์

4.2 ผลการวิเคราะห์องค์ประกอบเชิงยืนยัน

ผลการวิจัย การวิเคราะห์องค์ประกอบเชิงยืนยัน พบว่าปัจจัยที่มีอิทธิพลต่อสมรรถนะด้านเทคโนโลยีสารสนเทศของพนักงาน ปตท.จำกัด มีความสอดคล้องกับข้อมูลเชิงประจักษ์ที่ได้จากการเก็บข้อมูลจากกลุ่มตัวอย่างเป็นอย่างดี แสดงดังภาพที่ 8



ภาพที่ 8: การวิเคราะห์องค์ประกอบเชิงยืนยัน

ผู้วิจัยได้ทำการปรับโมเดล โดยพิจารณาจากความสอดคล้องทางทฤษฎีกับค่าดัชนีปรับแก้โมเดล (Modification Indices) ผล

พบว่า จากค่า Chi-square = 1094.713, df = 277, P-value = 0.194, RMSEA = 0.074 พบว่า การทดสอบโมเดลตามทฤษฎีมีความสอดคล้องกับข้อมูลเชิงประจักษ์ดี ค่าดัชนีวัดระดับความกลมกลืน (GFI) มีค่าเท่ากับ 0.932 ค่าดัชนีวัดระดับความกลมกลืนปรับแก้แล้ว (AGFI) มีค่าเท่ากับ 0.918 และค่าดัชนีรากของกำลังสองเฉลี่ยส่วนที่เกิน (RMR) มีค่าเท่ากับ 0.0059 จากการปรับแก้โมเดล แสดงให้เห็นว่า โมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์อย่างสมบูรณ์ และได้โมเดลรูปแบบความสัมพันธ์เชิงสาเหตุขององค์ประกอบวิเคราะห์ปัจจัยที่มีอิทธิพลต่อสมรรถนะด้านเทคโนโลยีสารสนเทศของพนักงาน ปตท.จำกัด โดยการวิเคราะห์องค์ประกอบเชิงยืนยัน

โดยค่าน้ำหนักขององค์ประกอบของตัวแปรทั้งหมดมีค่าเป็นบวกตั้งแต่ 0.486 - 0.806 และมีนัยสำคัญทางสถิติที่ระดับ 0.05 ทุกตัว โดยองค์ประกอบปัจจัยที่มีอิทธิพลต่อสมรรถนะด้านเทคโนโลยีสารสนเทศของพนักงาน ปตท.จำกัด ได้แก่ ด้านความรับผิดชอบต่อการใช้เทคโนโลยีสารสนเทศ ค่าน้ำหนักความสำคัญสูงสุด โดยมีค่าน้ำหนักความสำคัญเท่ากับ 17% รองลงมา ค่าน้ำหนักความสำคัญเท่ากับ 2 ค่า ด้านความสามารถพื้นฐานเกี่ยวกับการใช้เทคโนโลยีสารสนเทศ โดยมีค่าน้ำหนักความสำคัญเท่ากับ 16% ด้านความเชี่ยวชาญในการใช้เทคโนโลยีสารสนเทศ โดยมีค่าน้ำหนักความสำคัญเท่ากับ 16% ต่อมา ได้แก่ ด้านการใช้ประโยชน์ของเทคโนโลยีสารสนเทศ โดยมีค่าน้ำหนักความสำคัญเท่ากับ 15% ด้านความรู้ความเข้าใจเกี่ยวกับการใช้เทคโนโลยีสารสนเทศ มีค่าน้ำหนักความสำคัญเท่ากับ 13% ด้านการเห็นคุณค่าในเทคโนโลยีสารสนเทศ มีค่าน้ำหนักความสำคัญเท่ากับ 13% น้อยที่สุด ด้านการประยุกต์ใช้เทคโนโลยีสารสนเทศ ให้เกิดประโยชน์ มีค่าน้ำหนักความสำคัญเท่ากับ 10% แสดงดังตารางที่ 2

ตารางที่ 2: ผลการวิเคราะห์องค์ประกอบเชิงยืนยัน โมเดลการวัดองค์ประกอบ

องค์ประกอบ	น้ำหนักองค์ประกอบ	% น้ำหนักองค์ประกอบ
ด้านความรับผิดชอบต่อการใช้เทคโนโลยีสารสนเทศ	0.806	17
ด้านความสามารถพื้นฐานเกี่ยวกับการใช้เทคโนโลยีสารสนเทศ	0.760	16
ด้านความเชี่ยวชาญในการใช้เทคโนโลยีสารสนเทศ	0.749	16
ด้านการใช้ประโยชน์ของเทคโนโลยีสารสนเทศ	0.742	15
ด้านการเห็นคุณค่าในเทคโนโลยีสารสนเทศ	0.636	13
ด้านความรู้ความเข้าใจเกี่ยวกับการใช้เทคโนโลยีสารสนเทศ	0.627	13
ด้านการประยุกต์ใช้เทคโนโลยีสารสนเทศให้เกิดประโยชน์	0.486	10

5. สรุป

จากผลการวิจัย เรื่องการวิเคราะห์ปัจจัยที่มีอิทธิพลต่อสมรรถนะด้านเทคโนโลยีสารสนเทศของพนักงาน ปตท.จำกัด โดยการวิเคราะห์องค์ประกอบเชิงสำรวจ ในแต่ละด้าน ซึ่งมีทั้งหมด 7 ด้าน แล้วนำปัจจัยทั้งหมดมาทดสอบความสอดคล้องของโมเดลโดยใช้การวิเคราะห์องค์ประกอบเชิงยืนยัน ผลพบว่า โมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์เป็นอย่างดีและปัจจัยทั้ง 7 ด้านนี้ สามารถที่จะนำไปพัฒนาเพื่อสร้างหลักสูตรฝึกอบรมสมรรถนะด้านเทคโนโลยีสารสนเทศของพนักงาน ปตท.จำกัดในขั้นตอนต่อไปได้

6. เอกสารอ้างอิง

- [1] ไชยา เข็นแซ (2547) วิศวกรระดับ 7 แผนก เทคโนโลยีสารสนเทศ กองบำรุงรักษาระบบสื่อสาร ฝ่ายปฏิบัติการภาคกลาง
- [2] ฉัตรศิริ (2541-2543) การใช้ SPSS เพื่อการวิเคราะห์ข้อมูล, ความสัมพันธ์โครงสร้างเชิงเส้นเบื้องต้นสาขาการทดสอบและวัดผลการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ
- [3] นิพนธ์ ลือภักดินันท์ (2547: 71 - 72) , ปัจจัยที่มีผลต่อความสำเร็จขององค์การในการใช้ระบบสารสนเทศในการปฏิบัติงานของพนักงานบริษัท คาโอ อินดัสตรีเยล (ประเทศไทย) จำกัด บริหารธุรกิจ สาขาการจัดการ มหาวิทยาลัยศรีนครินทรวิโรฒ.
- [4] สุชาติ ประสิทธิ์รัฐสินธุ์. แบบจำลองสมการโครงสร้าง : การใช้โปรแกรม LISREL, PRELIS และ SIMPLIS ระเบียบวิธีการวิจัยทางสังคมศาสตร์.พิมพ์ครั้งที่ 12. กรุงเทพฯ : สถาบัน บัณฑิตพัฒนานบริหารศาสตร์,2546.
- [5] รัตติกรณ์ จงวิศาล การวิเคราะห์องค์ประกอบเชิงยืนยันของแบบวัดบรรยากาศ มหาวิทยาลัยศรีนครินทรวิโรฒ, 2544.
- [6] มนเชิธร รัตนศิริวงค์วุฒิ (2553) ปัจจัยที่มีอิทธิพลต่อสมรรถนะนักเทคโนโลยีสารสนเทศ โดยการวิเคราะห์องค์ประกอบเชิงยืนยันลำดับที่สองมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ.
- [7] มนเชิธร รัตนศิริวงค์วุฒิ (2552) ปัจจัยที่มีอิทธิพลต่อสมรรถนะนักเทคโนโลยีสารสนเทศ โดยใช้วิธีการวิเคราะห์องค์ประกอบเชิงสำรวจ. รายงานการประชุมสัมมนาทางวิชาการ มหาวิทยาลัยเทคโนโลยีราชมงคลตะวันออก คณะเทคโนโลยีสารสนเทศ
- [8] Li (2004), Chen, Li-Yuck. 2004. An Examination of The Relationships Among Leadership Behaviors, Knowledge Sharing, And Organizations' Marketing Effectiveness In Professional Service Firms That Have been Engaged in Strategic Alliances. Dissertation, Ph.D. (Business Administration). Graduate School The Nova Southeastern University. Photocopied. Retrieved May 16, 2005, from UMI ProQuest Digital Dissertation.

การจัดทำแนวทางการปฏิบัติในการรักษาความมั่นคงปลอดภัยของเทคโนโลยีสารสนเทศ ด้วยมาตรฐาน ISO/IEC 27001

กรณีศึกษา : สำนักงานรัฐบาลอิเล็กทรอนิกส์ (มหาชน) (สรอ.)

The Guidance for Security Operation of Information Technology using ISO/IEC 27001 Standard

A case study of Electronic Government Agency: EGA (Public Organization)

รัชชาภรณ์ สุภาพ (Ratchaporn Suphap)¹ และศักดิ์ชาย ตั้งวรรณวิทย์ (Sakchai Tangwannawit)²

¹ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

ratchaprons@hotmail.com¹, sakchait@kmutnb.ac.th²

บทคัดย่อ

งานวิจัยครั้งนี้มีวัตถุประสงค์เพื่อลดความเสี่ยงต่างๆ ในการทำงานด้านเทคโนโลยีสารสนเทศขององค์กร โดยการนำมาตรฐาน ISO/IEC 27001 มาประยุกต์ใช้โดยทำการศึกษาที่สำนักงานรัฐบาลอิเล็กทรอนิกส์ (องค์การมหาชน) ซึ่งมีวิธีการดำเนินงานแบ่งได้เป็น 4 ขั้นตอน 1) ศึกษามาตรฐานการจัดการความมั่นคงปลอดภัย ISO/IEC 27001 2) วิเคราะห์ความเสี่ยงด้านเทคโนโลยีสารสนเทศขององค์กร 3) พัฒนาแนวทางปฏิบัติในการรักษาความปลอดภัยทางด้านเทคโนโลยีสารสนเทศ ตามมาตรฐานด้านความปลอดภัยของข้อมูล ISO/IEC 27001 4) จัดทำข้อเสนอแนะเพื่อสร้างแนวทางปฏิบัติในการรักษาความมั่นคงปลอดภัยของเทคโนโลยีสารสนเทศโดยผู้วิจัยพัฒนาแบบสอบถาม ตามมาตรฐาน ISO/IEC 27001 ทั้ง 11 หมวด สำหรับกลุ่มตัวอย่างที่เป็นผู้ใช้งานเทคโนโลยีสารสนเทศในองค์กร การวิเคราะห์งานวิจัยนี้เป็นการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปร ผลการวิจัยพบว่าเจ้าหน้าที่ภายในองค์กรส่วนมากยังไม่ปฏิบัติตามมาตรฐานการรักษาความมั่นคงปลอดภัยสำหรับสารสนเทศขององค์กร ISO/IEC 27001

ผลสรุปจากการวิจัยพบว่าองค์กรควรมีการประเมินความเสี่ยงด้านเทคโนโลยีสารสนเทศ และจัดทำนโยบายความมั่นคงปลอดภัย

คำสำคัญ: การรักษาความปลอดภัยของเทคโนโลยีสารสนเทศ
มาตรฐาน ISO/IEC 27001

Abstract

The purpose of this study is to reduce the risks in information technology security applied in Electronic Government Agency (Public organization: EGA) using a standard ISO/IEC 27001. This work is divided into 4 steps, which are 1) Studying ISO/IEC 27001 2) Analyzing of information technology risk 3) Developing guidance for information technology as ISO/IEC 27001 and 4) Presenting recommend paper in security of information Technology. The researcher developed questionnaire which ISO/IEC 27001 of 11 categories standard. For evaluation, the sample group used information technology in EGA.

The results show that the most officers in the organization do not follow security standard. The organization should be assessment risk and develop the security policy for information technology.

Keyword: Security of Information Technology, Standard ISO/IEC 27001

1. บทนำ

ในยุคปัจจุบันรูปแบบการทำงานขององค์กรเกือบทุกแห่งทั่วโลกเกิดความเปลี่ยนแปลงจากเดิมไปอย่างมาก ด้วยความก้าวหน้าของเทคโนโลยีส่งผลให้คอมพิวเตอร์ อุปกรณ์เทคโนโลยีสารสนเทศ และระบบต่างๆ เข้ามามีบทบาทสำคัญ

ต่อกระบวนการทางธุรกิจอย่างหลีกเลี่ยงไม่ได้ ปรากฏการณ์เช่นนี้ส่งผลให้แนวโน้มในการดำเนินธุรกิจและความปลอดภัยขององค์กรจึงต้องพึ่งพาระบบเทคโนโลยีสารสนเทศ

ระบบสารสนเทศจึงมีความจำเป็นในการดำเนินงานของสำนักงานรัฐบาลอิเล็กทรอนิกส์ (องค์การมหาชน) (สรอ.) Electronic Government Agency : EGA ทั้งนี้ ระบบสารสนเทศอาจอยู่ได้ในหลายรูปแบบไม่ว่าจะเป็นเอกสาร สิ่งพิมพ์หรือถูกเก็บไว้ในสื่ออิเล็กทรอนิกส์ เพื่อเป็นการป้องกันเชิงรุกต่อความเสี่ยงที่อาจเกิดขึ้นจากภัยคุกคาม และรักษาความมั่นคงปลอดภัยให้กับทรัพย์สินของสำนักงาน โดยเฉพาะทรัพย์สินที่เป็นสารสนเทศสำหรับการดำเนินงานของสำนักงาน โดยการป้องกันให้พ้นจากภัยคุกคามทั้งจากภายในและภายนอก ทั้งที่เกิดจากความตั้งใจและไม่ตั้งใจก็ตาม รวมถึงเพื่อลดความเสียหายต่างๆ เพื่อรักษาไว้ซึ่งความสามารถในการดำเนินงานได้อย่างต่อเนื่องความเสี่ยงที่มีผลต่อความมั่นคงปลอดภัยของข้อมูล รวมไปถึงระบบสารสนเทศภายในองค์กร นอกจากนี้สิ่งที่กล่าวไปแล้วข้างต้นนั้น ยังมีสาเหตุมาจากภัยธรรมชาติ การจลาจล อุบัติเหตุ เหตุการณ์ร้ายแรงต่างๆ ที่ไม่คาดคิดสามารถเกิดขึ้นได้ตลอดเวลา ซึ่งบางเหตุการณ์ก็อยู่เหนือการควบคุม และการป้องกันขององค์กร เพื่อให้องค์กรสามารถฝ่าวิกฤตที่เกิดขึ้นและสามารถกลับมาดำเนินธุรกิจได้อีกครั้ง

ดังนั้น จึงจำเป็นต้องมีแนวทางการปฏิบัติในการรักษาความมั่นคงปลอดภัยทางด้านเทคโนโลยีสารสนเทศ ตามมาตรฐานด้านความปลอดภัยของข้อมูล ISO/IEC 27001 เพื่อกำหนดเป็นแนวทางและวิธีปฏิบัติให้กับองค์กร เมื่อองค์กรนำแนวทางการปฏิบัติในการรักษาความมั่นคงปลอดภัยทางด้านเทคโนโลยีสารสนเทศ ตามมาตรฐานด้านความปลอดภัยของข้อมูล ISO/IEC 27001 ไปปฏิบัติอย่างจริงจังนั้น จะทำให้องค์กรปฏิบัติงานอย่างมีประสิทธิภาพและดำเนินธุรกิจอย่างต่อเนื่อง รวมทั้งส่งผลให้พนักงานทั้งหมดในองค์กรตื่นตัวและตระหนักถึงความปลอดภัยที่เกิดขึ้นได้ ซึ่งจะทำให้องค์กรมีระบบความปลอดภัยในการปฏิบัติงานจากพนักงาน และทำให้เกิดความปลอดภัยกับข้อมูลสารสนเทศที่สำคัญที่สำคัญขององค์กร

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 มาตรฐานด้านความปลอดภัยของข้อมูลมาตรฐาน ISO/IEC 27001 (Information security management system)

มาตรฐานที่เกี่ยวกับการจัดการในเรื่องความปลอดภัยของข้อมูล ได้รับการพัฒนามาจาก Information Security Management Standard BS7799 [1] ออกโดย British Standard Institute (BSI) ซึ่งมาตรฐานนี้เป็นส่วนหนึ่งของมาตรฐาน ISO (International Standard Organization) ที่ถูกกำหนดเพื่อเป็นแนวทางการจัดการด้านความปลอดภัยของข้อมูลภายในองค์กร โดยมีการกำหนดแนวทางสำหรับองค์กรในด้านการปฏิบัติงาน การจัดการเพื่อให้เกิดประสิทธิภาพ การพัฒนามาตรฐานความปลอดภัย เพื่อการสร้างความมั่นใจในการเชื่อมโยงระหว่างองค์กร เนื่องจากข้อมูลขององค์กรถือว่าเป็นทรัพย์สินที่มีความสำคัญ ดังนั้นการรักษาความปลอดภัยของข้อมูล, การวิเคราะห์และการบริหารความเสี่ยงที่อาจเกิดขึ้นจึงถือเป็นสิ่งที่สำคัญในการบริหารงานองค์กรให้มีประสิทธิภาพ

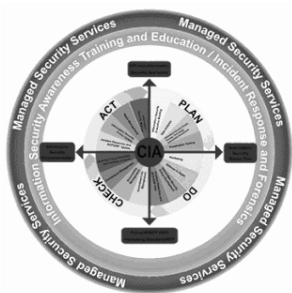
มาตรฐาน ISO/IEC 27001 จะประกอบไปด้วย 11 หมวดดังต่อไปนี้

หมวดที่ 1: นโยบายความมั่นคงปลอดภัย (Security policy),
 หมวดที่ 2: โครงสร้างด้านความมั่นคงปลอดภัยสำหรับองค์กร (Organization of Information security),
 หมวดที่ 3: การบริหารจัดการทรัพย์สินขององค์กร (Asset management),
 หมวดที่ 4: ความมั่นคงปลอดภัยที่เกี่ยวข้องกับบุคลากร (Human resources security),
 หมวดที่ 5: การสร้างความมั่นคงปลอดภัยทางกายภาพและสิ่งแวดล้อม (Physical and environmental security),
 หมวดที่ 6: การบริหารจัดการด้านการสื่อสารและการดำเนินงานของเครือข่ายสารสนเทศขององค์กร (Communications and operations management),
 หมวดที่ 7: การควบคุมการเข้าถึง (Access control),
 หมวดที่ 8: การจัดการพัฒนา และการบำรุงรักษาระบบสารสนเทศ (Information system acquisition, development and maintenance),
 หมวด 9: การบริหารจัดการเหตุการณ์ที่เกี่ยวข้องกับความมั่นคงปลอดภัยขององค์กร (Information security incident management),
 หมวด 10: การบริหารความต่อเนื่องในการดำเนินงานขององค์กร (Business continuity management) และหมวดที่ 11: การปฏิบัติตามข้อกำหนด (Compliance)

2.2 ทฤษฎีที่เกี่ยวข้องกับการรักษาความมั่นคงปลอดภัย (Information Security)

ความปลอดภัย (Security) [2] หมายถึง “ความรู้สึกรู้สึกหรือสถานะที่ปลอดภัย-การปราศจากอันตรายใด” ในมุมมองของผู้บริหารแต่ละท่านความปลอดภัยต้องเกิดจากการวางแผน, การดำเนินงาน, บุคลากร การสั่งการ และการควบคุมที่เหมาะสมและองค์ประกอบหลักๆ ที่ได้รับการยอมรับอย่างกว้างขวางของความปลอดภัยของข้อมูล

โมเดล ISMF version 2 ที่พัฒนาขึ้นจึงได้ยึดแนวความคิดด้าน Risk Management เป็นหลัก ซึ่งสามารถช่วยให้องค์กรสามารถดำเนินการด้านการรักษาความปลอดภัยของข้อมูลได้อย่างเหมาะสม โดยคำนึงถึง Risk Tolerance และ Risk Acceptance เป็นหลัก ซึ่งจะช่วยให้องค์กรได้รับผลตอบแทนจากการลงทุน(ROI) จากโครงการรักษาความปลอดภัยของข้อมูลได้ในระดับสูงสุดประกอบไปด้วย 4 งานหลัก (Plan, Do, Check, Act) และแต่ละงานหลักมี 4 งานย่อย (รูปทรงกราฟวงกลมที่อยู่ถัดจากชื่องานหลัก) รวมทั้งหมด 16 งานย่อย โดยแต่ละงานหลักนั้นจะมีลูกศรสีแดงชี้ออกมาซึ่งหมายถึงผลลัพธ์สุดท้ายของแต่ละงานหลัก และองค์ประกอบสุดท้ายคือวงแหวนรอบนอกซึ่งแสดงถึงงานที่จะต้องทำอย่างต่อเนื่องสม่ำเสมอตลอดโครงการไม่จำกัดว่าจะต้องทำในช่วงงานหลักใด สำหรับความหมายและหน้าที่ขององค์ประกอบต่างๆ ดังรูปภาพที่ 1



ภาพที่ 1: โมเดล Information Security Management Framework

[www.acisonline.net]

2.3 งานวิจัยที่เกี่ยวข้อง

ทรงชล มหารมณ [3] ได้ทำการวิจัยเรื่อง การวิเคราะห์ความเสี่ยงการรักษาความปลอดภัยของข้อมูลสำหรับองค์กรขนาด

ใหญ่และขนาดกลาง ในกรุงเทพมหานคร จากการศึกษาภัยคุกคาม และปัจจัยที่ก่อให้เกิดความเสี่ยงที่เป็นพฤติกรรม และลักษณะการทำงานของบุคลากรในองค์กร ในการรักษาความปลอดภัยข้อมูลสารสนเทศขององค์กรขนาดกลาง และขนาดใหญ่ ได้แก่ ความเชื่อถือ (Confidentiality) ความถูกต้องและสมบูรณ์ (Integrity) และความสามารถในการนำมาใช้ประโยชน์ได้เมื่อต้องการ (Availability) ความเสี่ยงที่เป็นไปได้อาจเกิดขึ้นกับความปลอดภัยของข้อมูล

ภูมิินทร์ ภูดวงสี [4] ได้ทำการวิจัยเรื่องการศึกษาแนวทางการพัฒนานโยบายความมั่นคงปลอดภัยในองค์กร ISO 27001 กรณีศึกษาบริษัท NEC Corporation (Thailand) Ltd. เพื่อนำมาใช้และเพื่อลดความเสี่ยงที่จะเกิดขึ้นกับระบบสารสนเทศขององค์กร อีกทั้งเพื่อกำหนดเป็นกฎข้อบังคับให้กับพนักงาน และสร้างความตระหนักถึงความปลอดภัยในการใช้สารสนเทศเพื่อประสิทธิภาพในการทำงานและสร้างความเชื่อมั่นให้กับลูกค้าในการใช้ผลิตภัณฑ์ของบริษัทโดยนำมามาตรฐานความมั่นคงปลอดภัยมาประยุกต์ปรับใช้ให้เหมาะสมกับองค์กร

เพ็ญประภา พิพัฒนาโฆษิต [5] ได้ทำการวิจัยเรื่องการสร้างมาตรฐานการบริหารความปลอดภัยของข้อมูล โดยการประยุกต์ใช้ ISO 17799 กรณีศึกษา ของบริษัทให้บริการด้านการจัดซื้อออนไลน์เพื่อทำการศึกษาแนวทางขั้นตอนการดำเนินงาน สำหรับการประยุกต์ใช้มาตรฐาน ISO 17799 ในการรักษาความปลอดภัยของข้อมูล ระบบคอมพิวเตอร์และระบบสารสนเทศในองค์กรที่ให้บริการด้านการจัดซื้อออนไลน์ (e-Procurement) และเพื่อศึกษาปัจจัยที่มีผลต่อความเสี่ยงด้านความปลอดภัยของข้อมูลที่อาจเกิดขึ้นกรณีองค์กรขาดการป้องกัน และสรุปเป็นนโยบายนำเสนอแก่ผู้บริหารขององค์กรสำหรับนำไปจัดทำนโยบายเพิ่มเติมและปรับปรุงองค์กรเพื่อให้เหมาะสมกับการพัฒนาทางธุรกิจ

นายนิรทธี จินดาจามร, นางสาววิญดา วรรณโกสิน, นายอรรถพล คุ่มเศรณี [6] ได้ทำการวิจัยเรื่องการศึกษาและจัดทำแนวปฏิบัติในการรักษาความมั่นคงปลอดภัยด้านสารสนเทศตามพระราชบัญญัติว่าด้วยธุรกรรมทางอิเล็กทรอนิกส์ พ.ศ. 2544 เพื่อกำหนดแนวปฏิบัติที่เป็นรูปธรรมในการรักษาความมั่นคงปลอดภัยด้านสารสนเทศให้กับองค์กรภาครัฐและภาคเอกชนที่เกี่ยวข้องกับการประกอบธุรกรรมทาง

อิเล็กทรอนิกส์ อ้างอิงตามมาตรฐานความปลอดภัยสากล ISO/IEC 17799:2005 ประกอบร่างพระราชกฤษฎีกา มาตรา 25, มาตรา 32 และมาตรา 35 แห่งพระราชบัญญัติว่าด้วยธุรกรรมทางอิเล็กทรอนิกส์ พ.ศ. 2544

กิตติ วัฒนกิจรุ่งโรจน์ และณรงค์ศักดิ์ สุขเกื้อ [7] ได้ทำการวิจัยเรื่องแนวทางในการรักษาความปลอดภัยข้อมูล โดยการจัดทำนโยบายการรักษาความปลอดภัยของระบบคอมพิวเตอร์ และเครือข่าย ของคณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ ซึ่งมีการกำหนดระเบียบการใช้งานเครือข่าย และการสร้างความปลอดภัยให้กับระบบ เช่น การออกแบบโครงสร้างของระบบเครือข่ายที่มีความปลอดภัย การติดตั้งระบบตรวจจับผู้บุกรุกทางเครือข่ายและการประเมินระบบรักษาความปลอดภัยข้อมูล เป็นต้น พร้อมทั้งได้ทำสรุปผลการประเมิน และเสนอแนวทางในการปรับปรุงระบบรักษาความปลอดภัยให้ดียิ่งขึ้น

จากการศึกษางานวิจัยที่เกี่ยวข้องทำให้ผู้วิจัยพบว่าปัจจุบันนี้คือยุคของการนำคอมพิวเตอร์มาใช้ในการดำเนินงานอย่างแพร่หลายทั่วทั้งองค์กร ซึ่งองค์กรส่วนมากจำเป็นต้องใช้ระบบอินเทอร์เน็ตในการดำเนินงาน แลกเปลี่ยนข้อมูล และติดต่อสื่อสารต่างๆ รวมถึงตัวบุคคลมีความสำคัญที่องค์กรจำเป็นต้องป้องกันเพราะสภาพแวดล้อมในปัจจุบันที่ต้องติดต่อสื่อสารกันมากขึ้นทำให้เกิดปัญหาและจุดอ่อนหลายรูปแบบที่ทำให้สารสนเทศนั้นเสียหายได้

ดังนั้นผู้วิจัยจึงเห็นว่ามาตรฐานและหลักปฏิบัติในการป้องกันควบคุมข้อมูลก็คือมาตรฐานการรักษาความปลอดภัย ISO/IEC 27001 เพื่อเป็นการสร้างความมั่นใจว่าข้อมูลและสารสนเทศความลับยังถูกเก็บรักษาอยู่ครบถ้วนปลอดภัยเป็นอย่างดี

3. วิธีการดำเนินการวิจัย

ในการวิจัยครั้งนี้ ซึ่งผู้วิจัยได้แบ่งขั้นตอนวิธีการดำเนินงานไว้ โดยมีรายละเอียดดังต่อไปนี้

3.1 ศึกษาจากทฤษฎีที่เกี่ยวข้อง โดยทำการศึกษานโยบายการจัดการความมั่นคงปลอดภัยสารสนเทศขององค์กร ซึ่งรวมถึงมาตรฐานการจัดการความมั่นคงปลอดภัย ISO/IEC 27001

3.2 วิเคราะห์ความเสี่ยงจากภัยคุกคาม และช่องโหว่ด้านเทคโนโลยีสารสนเทศขององค์กร

3.3 จัดทำแบบสอบถามพนักงานในองค์กร เพื่อสำรวจความคิดเห็นของผู้ใช้งานที่มีต่อเทคโนโลยีสารสนเทศขององค์กร ผู้วิจัยจึงได้ใช้วิธีการสำรวจข้อมูล โดยใช้แบบสอบถาม (Questionnaire) แก่บุคลากรขององค์กร จำนวน 133 คน เพื่อทำการสำรวจความคิดเห็นที่มีต่อการดำเนินงานทางด้านระบบสารสนเทศขององค์กร แต่ในการแจกแบบสอบถามจริงได้แจกไปทั้งหมด 150 ชุด เพื่อป้องกันการสูญหาย ผลปรากฏได้รับแบบสอบถามกลับมาทั้งสิ้นจำนวน 133 ชุด คิดเป็น 88.67% ของจำนวนแบบสอบถามทั้งหมดที่แจก

3.4 นำผลที่ได้มาวิเคราะห์ทางสถิติ เพื่อจัดทำแนวทางปฏิบัติในการรักษาความมั่นคงปลอดภัยของเทคโนโลยีสารสนเทศตามมาตรฐาน ISO/IEC 27001

3.4.1 การเลือกกลุ่มตัวอย่างข้อมูลที่นำมาวิเคราะห์ จะเป็นข้อมูลที่ได้จากการตอบแบบสอบถาม โดยกำหนดขนาดของกลุ่มตัวอย่าง โดยประมาณค่าสัดส่วนของประชากร (Proportions sample size determination) คำนวณจากสูตรทาร์ยามาน (Taro Yamane, 1973:125) ซึ่งมีสูตรการคำนวณดังนี้

$$n = \frac{N}{1 + Ne^2}$$

3.4.2 การวิเคราะห์ข้อมูลในการวิจัยนี้มุ่งใช้สถิติ โดยใช้โปรแกรม SPSS เพื่อหาค่าร้อยละ ค่าเฉลี่ย ค่าความเป็นอิสระต่อกันของสองประชากร (Chi-square test) และการวิเคราะห์ผลแบบ One Sample T-test ในการวิเคราะห์ความคิดเห็นที่ได้จากกลุ่มตัวอย่างที่เป็นผู้ใช้งานเทคโนโลยีสารสนเทศ

3.5 นำผลที่ได้จากการวิเคราะห์ทางสถิติ จัดทำเป็นแนวทางปฏิบัติในการรักษาความมั่นคงปลอดภัยด้านเทคโนโลยีสารสนเทศและการสื่อสารที่เหมาะสมกับองค์กร เพื่อให้เกิดประโยชน์และสามารถนำไปใช้ปฏิบัติได้จริง

3.6 สรุปวิเคราะห์ผลการศึกษาและจัดทำข้อเสนอแนะที่เหมาะสมกับองค์กรต่อไป

4. ผลการดำเนินงาน

ผลการวิเคราะห์ข้อมูลการแสดงความคิดเห็นของผู้ใช้เทคโนโลยีสารสนเทศเป็นข้อมูลที่ได้จากการตอบแบบสอบถามของกลุ่มตัวอย่าง สรุปได้ดังนี้ 1) คิดเป็นเพศชายเท่ากับ 51.88% และเพศหญิง เท่ากับ 48.12% 2) อายุ 20-30 ปี

เท่ากับ 45.11%, อายุ 31-40 เท่ากับ 50.38%, 41-50 ปี เท่ากับ 4.51% และ มากกว่า 51 ปี เท่ากับ 0% 3) วุฒิการศึกษาปริญญาตรี เท่ากับ 68.42%, ปริญญาโท เท่ากับ 31.58%, ปริญญาเอก เท่ากับ 0% 4) ประสบการณ์การทำงานในปัจจุบัน น้อยกว่า 1 ปี เท่ากับ 11.28%, 1-2 ปี เท่ากับ 27.07%, 3-4 ปี เท่ากับ 36.09%, 5-10 ปี เท่ากับ 22.56 และ มากกว่า 10 ปี เท่ากับ 3.01% และ 5) แบ่งตามตำแหน่งงาน ระดับปฏิบัติการ เท่ากับ 97.74% และ ระดับหัวหน้าส่วนงาน/ผู้จัดการ เท่ากับ 2.26%

จากผลการวิเคราะห์โดยการหาค่า Chi-Square และ One Sample T-test พบว่า

4.1 กำหนดสมมติฐาน

เจ้าหน้าที่สำนักงานปฏิบัติตามมาตรฐาน ISO/IEC 27001 เท่ากับระดับดี $H_0 \mu = 4.00$

เจ้าหน้าที่สำนักงานปฏิบัติตามมาตรฐาน ISO/IEC 27001 ไม่เท่ากับระดับดี $H_1 \mu \neq 4.00$

4.2 กำหนดระดับนัยสำคัญ

แอลฟา $\alpha = 0.05$

4.3 สรุปผล ยอมรับหรือปฏิเสธ H_0 โดยจะปฏิเสธ H_0 เมื่อ

$sig < \text{แอลฟา } \alpha$

ตารางที่ 1: ผลการศึกษาความสัมพันธ์ของตัวแปร

ตัวแปร	Pearson Chi-Square	ผลการวิจัย
เพศและสถานภาพ	0.784	ไม่มีความสัมพันธ์กัน
เพศและอายุ	0.014	มีความสัมพันธ์กัน
เพศและการศึกษา	0.298	ไม่มีความสัมพันธ์กัน
เพศและประสบการณ์	0.444	ไม่มีความสัมพันธ์กัน
เพศและตำแหน่งงาน	0.092	ไม่มีความสัมพันธ์กัน
สถานภาพและอายุ	0.006	มีความสัมพันธ์กัน
สถานภาพและการศึกษา	0.215	ไม่มีความสัมพันธ์กัน
สถานภาพและประสบการณ์ทำงาน	0.045	มีความสัมพันธ์กัน
สถานภาพและตำแหน่งงาน	0.006	มีความสัมพันธ์
อายุและการศึกษา	0.071	ไม่มีความสัมพันธ์
อายุและประสบการณ์ทำงาน	0.000	มีความสัมพันธ์
อายุและตำแหน่งงาน	0.221	ไม่มีความสัมพันธ์
การศึกษาและประสบการณ์ทำงาน	0.032	มีความสัมพันธ์
การศึกษาและตำแหน่งงาน	0.186	ไม่มีความสัมพันธ์
ตำแหน่งงานและประสบการณ์	0.032	มีความสัมพันธ์

ตารางที่ 2: ผลการวิจัยของกลุ่มตัวอย่างที่มีต่อการดำเนินงานด้านเทคโนโลยีสารสนเทศและการสื่อสารขององค์กร

มาตรฐาน ISO/IEC 27001	ระดับนัยสำคัญ	ผลการวิจัย
หมวดที่ 1: นโยบายความมั่นคงปลอดภัย	0.023	เจ้าหน้าที่สำนักงานรัฐบาลอิเล็กทรอนิกส์ส่วนมากยังไม่ปฏิบัติตามมาตรการด้านนโยบายความมั่นคงปลอดภัยสารสนเทศขององค์กรจึงยังไม่อยู่ในระดับดี
หมวดที่ 2: โครงสร้างทางด้านการความมั่นคงปลอดภัยสำหรับองค์กร	0.040	เจ้าหน้าที่สำนักงานรัฐบาลอิเล็กทรอนิกส์ส่วนมากยังไม่ปฏิบัติตามมาตรการด้านโครงสร้างทางด้านการความมั่นคงปลอดภัยสำหรับองค์กรจึงยังไม่อยู่ในระดับดี
หมวดที่ 3: การบริหารจัดการทรัพย์สินขององค์กร	0.081	เจ้าหน้าที่สำนักงานรัฐบาลอิเล็กทรอนิกส์ปฏิบัติตามมาตรการด้านหน้าที่ความรับผิดชอบต่อทรัพย์สินขององค์กรอยู่ในระดับดี
หมวดที่ 4: ความมั่นคงปลอดภัยที่เกี่ยวข้องกับบุคลากร	0.000	เจ้าหน้าที่สำนักงานรัฐบาลอิเล็กทรอนิกส์ส่วนมากยังไม่ปฏิบัติตามมาตรการด้านความมั่นคงปลอดภัยที่เกี่ยวข้องกับบุคลากรจึงยังไม่อยู่ในระดับดี
หมวดที่ 5: การสร้างความมั่นคงปลอดภัยทางกายภาพและสิ่งแวดล้อม	0.000	เจ้าหน้าที่สำนักงานรัฐบาลอิเล็กทรอนิกส์ส่วนมากยังไม่ปฏิบัติตามมาตรการด้านการสร้างความมั่นคงปลอดภัยทางกายภาพและสิ่งแวดล้อมจึงยังไม่อยู่ในระดับดี
หมวดที่ 6: การบริหารจัดการด้านการสื่อสารและการดำเนินงานของเครือข่าย	0.000	เจ้าหน้าที่สำนักงานรัฐบาลอิเล็กทรอนิกส์ส่วนมากยังไม่ปฏิบัติตามมาตรการด้านการบริหารจัดการด้านการสื่อสารและการดำเนินงานของเครือข่าย
หมวดที่ 7: การควบคุมการเข้าถึง	0.024	เจ้าหน้าที่สำนักงานรัฐบาลอิเล็กทรอนิกส์ส่วนมากยังไม่ปฏิบัติตามมาตรการด้านการจัดการการเข้าถึงระบบของพนักงานจึงยังไม่อยู่ในระดับดี
หมวดที่ 8: การจัดหา การพัฒนา และการบำรุงรักษาระบบสารสนเทศ	0.005	เจ้าหน้าที่สำนักงานรัฐบาลอิเล็กทรอนิกส์ส่วนมากยังไม่ปฏิบัติตามมาตรการด้านการจัดหา การพัฒนา และการบำรุงรักษาระบบสารสนเทศจึงยังไม่อยู่ในระดับดี

ตารางที่ 2: ผลการวิจัยของกลุ่มตัวอย่างที่มีต่อการดำเนินงานด้าน

เทคโนโลยีสารสนเทศและการสื่อสารขององค์กร (ต่อ)

มาตรฐาน ISO/IEC 27001	ระดับนัยสำคัญ	ผลการวิจัย
หมวดที่ 9: การบริหารจัดการเหตุการณ์ที่เกี่ยวข้องกับความมั่นคงปลอดภัยขององค์กร	0.002	เจ้าหน้าที่สำนักงานรัฐบาลอิเล็กทรอนิกส์ ส่วนมากยังไม่ปฏิบัติตามมาตรการด้านการบริหารจัดการเหตุการณ์ที่เกี่ยวข้องกับความมั่นคงปลอดภัยสารสนเทศจึงยังไม่อยู่ในระดับดี
หมวดที่ 10: การบริหารความต่อเนื่องในการดำเนินงานขององค์กร	0.078	เจ้าหน้าที่สำนักงานรัฐบาลอิเล็กทรอนิกส์ ปฏิบัติตามมาตรการด้านการบริหารความต่อเนื่องของการดำเนินงานขององค์กรอยู่ในระดับดี
หมวดที่ 11: การปฏิบัติตามข้อกำหนด	0.004	เจ้าหน้าที่สำนักงานรัฐบาลอิเล็กทรอนิกส์ ส่วนมากยังไม่ปฏิบัติตามมาตรการด้านการปฏิบัติตามข้อกำหนดทางด้านกฎหมายหรือนโยบาย และบทลงโทษจึงยังไม่อยู่ในระดับดี

5. สรุป อภิปรายผลและข้อเสนอแนะ

จากผลการการวิจัย ในปัจจุบันสามารถสรุปผลได้ดังนี้ หากองค์กรต้องการที่ก้าวสู่การรับรองการดำเนินงานทางด้านสารสนเทศตามมาตรฐาน ISO/IEC 27001 ทั้ง 11 หมวดองค์กรฯจึง ควรให้ความสำคัญและปฏิบัติตามในส่วนของหมวดที่ยังไม่ได้ดำเนินงานซึ่งมีเปอร์เซ็นต์มากถึง 81.81%

ซึ่งแนวทางการปฏิบัติในการรักษาความมั่นคงปลอดภัยของเทคโนโลยีสารสนเทศ ด้วยมาตรฐาน ISO/IEC 27001 จะทำให้ทราบถึงปัญหา แนวทางและวิธีปฏิบัติต่างๆ ซึ่งในบางข้อการควบคุมที่องค์กรยังไม่มีการบริหารจัดการที่ดี จึงนำผลการวิจัยที่พบเหล่านี้มาจัดทำเป็นแนวทางการปฏิบัติในการรักษาความปลอดภัยของเทคโนโลยีสารสนเทศของ ตามมาตรฐาน ISO/IEC 27001

ข้อเสนอแนะการวิจัยครั้งต่อไปองค์กรควรจัดทำแนวทางการปฏิบัติในการรักษาความปลอดภัยของเทคโนโลยีสารสนเทศในระดับเบื้องต้น และเมื่อองค์กรมีความพร้อมในระดับเบื้องต้นแล้วควรพัฒนาอย่างต่อเนื่องในระดับกลางและระดับสูงต่อไป

ซึ่งในการจัดทำแบบประเมิน และแบบสอบถามสำหรับสำรวจความคิดเห็น ต้องเข้าใจว่าพนักงานในองค์กรมีความรู้และความเข้าใจแต่ละเรื่องที่แตกต่างกัน ดังนั้นต้องทำความเข้าใจพร้อมทั้งอธิบายความหมายของคำถามให้ชัดเจนก่อน หรือควรปรับปรุงคำถามให้ง่ายต่อความเข้าใจ เพื่อให้ได้ผลจากการตอบแบบสอบถามเป็นผลที่สามารถนำไปวิเคราะห์ได้จริง และเพื่อเวลานำมาตรฐานกลับมาประยุกต์ใช้ ได้จัดความเข้มข้นในการใส่ใจกับแต่ละกลุ่มตัวอย่างที่เหมาะสม ซึ่งเป็นการเตรียมความพร้อมและแก้ไขให้ตรงจุดและลดระยะเวลาในการดำเนินการ

เอกสารอ้างอิง

- [1] ทวีศักดิ์ กอนันตกุลชฎา มาศ ชูระเศรษฐ, พันธศักดิ์ ศิริรัชตพงษ์, กุสุรางคณา วายุภาพ, ศิวรักษ์ ศิวโฆภยธรรม. “มาตรฐานการรักษาความมั่นคงปลอดภัยในการประกอบธุรกรรมทางอิเล็กทรอนิกส์”, ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ สำนักพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ กระทรวงวิทยาศาสตร์และเทคโนโลยี, 2550.
- [2] ปริญา หอมเอนก, “เรียนรู้โมเดลในการบริหารจัดการรักษาความปลอดภัยของข้อมูลอย่างเป็นระบบและมีประสิทธิภาพ”, แหล่งที่มา : <http://www.acisonline.net>.
- [3] ทรงชล มหารมณ, “การวิเคราะห์ความเสี่ยงการรักษาความปลอดภัยของข้อมูลสำหรับองค์กรขนาดใหญ่และขนาดกลาง ในกรุงเทพมหานคร”, วิทยาลัยนวัตกรรมการอุดมศึกษา มหาวิทยาลัยธรรมศาสตร์, 2548.
- [4] กุมินทร์ กุศลวงษ์, “การศึกษาแนวทางการพัฒนานโยบายความมั่นคงปลอดภัยในองค์กร ISO 27001 กรณีศึกษาบริษัท NEC Corporation (Thailand) Ltd. จำกัด”, วิทยาลัยนวัตกรรมการอุดมศึกษา มหาวิทยาลัยธรรมศาสตร์, 2550.
- [5] เพ็ญประภา พิพัฒนาโมชนิด, “การสร้างมาตรฐานการบริหารความปลอดภัยของข้อมูลโดยการประยุกต์ใช้ ISO 17799”กรณีศึกษาของบริษัทให้บริการด้านการจัดซื้อออนไลน์”, วิทยาลัยนวัตกรรมการอุดมศึกษา มหาวิทยาลัยธรรมศาสตร์, 2552.
- [6] นายนิธิร์ จินดาจามร, นางสาววิชุดา วรรณ โภคิน, นายอรุณพล คุ่มเสริม, “การศึกษาและจัดทำแนวปฏิบัติในการรักษาความมั่นคงปลอดภัยด้านสารสนเทศ ตามพระราชบัญญัติว่าด้วยธุรกรรมทางอิเล็กทรอนิกส์” พ.ศ. 2544”, มหาวิทยาลัยธรรมศาสตร์ 2550.
- [7] กิตติ วัฒนกิจรุ่งโรจน์, ณรงค์ศักดิ์ สุขเกื้อ, “แนวทางในการรักษาความปลอดภัยข้อมูล”, คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2552.

การประยุกต์ใช้มาตรฐานไอทีลกับระบบสนับสนุนการให้บริการอินเทอร์เน็ต An Application of ITIL Standard for Internet Service Support System

เบญจพร ปรีพินิจชัย (Benjaporn Paripinidchai)¹ และมหศักดิ์ เกตุน้ำ (Mahasak Ketcham)²

^{1,2} ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
paripinidchai@hotmail.com¹, maoquee@hotmail.com²

บทคัดย่อ

งานวิจัยนี้เป็นการนำเสนอการประยุกต์ใช้มาตรฐานไอทีลกับระบบสนับสนุนการให้บริการอินเทอร์เน็ต มีวัตถุประสงค์เพื่อศึกษาการนำมาตรฐานไอทีลมาประยุกต์ใช้กับการให้บริการอินเทอร์เน็ต เพื่อลดเวลาและขั้นตอนการทำงาน รวมไปถึงเพิ่มประสิทธิภาพในการให้บริการอินเทอร์เน็ต และเพิ่มความพึงพอใจในการใช้บริการจากผู้ให้บริการ โดยนำกระบวนการจัดการระดับบริการ การจัดการเหตุการณ์ไม่ปกติ และการจัดการปัญหาตามมาตรฐานไอทีลมาประยุกต์ใช้โดยมีลักษณะการทำงานในรูปแบบเว็บแอปพลิเคชัน เมื่อทำการประเมินประสิทธิภาพของระบบ พบว่า มีค่าเฉลี่ยเท่ากับ 4.20 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.51 สรุปได้ว่า ประสิทธิภาพของการประยุกต์ใช้มาตรฐานไอทีลกับระบบสนับสนุนการให้บริการอินเทอร์เน็ตอยู่ในระดับดี ช่วยทำให้สามารถจัดการกับเหตุการณ์ไม่ปกติเป็นไปตามข้อตกลงระดับบริการ ทั้งยังมีการนำเหตุการณ์ไม่ปกติมาวิเคราะห์ถึงสาเหตุที่แท้จริง เพื่อป้องกันไม่ให้ปัญหานั้นเกิดขึ้นอีก ส่งผลให้มีประสิทธิภาพในการให้บริการ

คำสำคัญ: ไอทีล การจัดการระดับบริการ การจัดการเหตุการณ์ไม่ปกติ การจัดการปัญหา

Abstract

This research presents an application of ITIL standard for internet service support system. The purpose of this study is to apply ITIL standard to internet service support system to decrease time and working processes. In addition to increase Internet services efficiency and customer satisfaction. By using service level

management, incident management and problem management apply based on ITIL standard via web application form. The overall result can be improved service efficiency. The system's efficiency results found mean equaled 4.20 and standard deviation equaled 0.51 consecutively, thus grading the system efficiency at good level. It will help to handle with anomaly incident to meet service level agreements (SLAs), additionally it will also analyze the incident to find out the root cause and can prevent from recurrence.

Keywords: ITIL, Service Level Management, Incident Management, Problem Management

1. บทนำ

ปัจจุบันเทคโนโลยีสารสนเทศเข้ามาเป็นส่วนหนึ่งในการสนับสนุนการดำเนินกิจกรรมทางธุรกิจ โดยเฉพาะภาคธุรกิจโทรคมนาคมที่มีอัตราการเติบโตและแข่งขันกันอย่างสูง จึงต้องมีกระบวนการในการดำเนินงานอย่างเป็นขั้นตอน เพื่อเพิ่มประสิทธิภาพในการให้บริการและทำให้องค์กรดำเนินธุรกิจบรรลุเป้าหมาย จึงได้มีการนำกระบวนการที่เป็นมาตรฐานในการบริหารการให้บริการมาเป็นเครื่องมือในการบริหารจัดการข้อมูล แก่ใจเหตุการณ์ไม่ปกติจากการให้บริการอินเทอร์เน็ตของผู้ใช้บริการให้กลับมาใช้งานได้ปกติโดยเร็วที่สุด ซึ่งจะส่งผลให้มีการพัฒนาและปรับปรุงการปฏิบัติงานให้มีคุณภาพและประสิทธิภาพมากขึ้น ทั้งยังทำให้ระบบการดำเนินงานเป็นไปในทิศทางที่ดีขึ้นอีกด้วย

ปัญหาคุณภาพและมาตรฐานการให้บริการ ส่งผลต่อการดำเนินธุรกิจขององค์กร หลายองค์กรจึงได้นำมาตรฐานต่าง ๆ มาใช้ในการพัฒนาคุณภาพการให้บริการสารสนเทศและรักษา

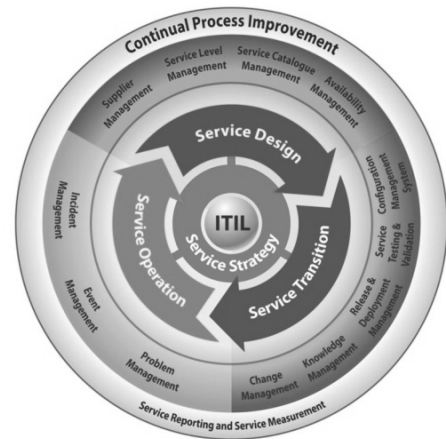
มาตรฐานการให้บริการแก่ผู้ใช้บริการ ซึ่งในปัจจุบันได้มีการนำมาตรฐานมากมายเข้ามาประยุกต์ใช้กับการบริหารจัดการด้านเทคโนโลยีสารสนเทศภายในองค์กร เช่น ไอทิล, COBIT และ ISO/IEC 27002 [1] พบว่า ควรใช้อิทิลในการกำหนดกลยุทธ์วางแผน และกระบวนการในการดำเนินงานการให้บริการสารสนเทศ ใช้มาตรฐาน COBIT ในการชี้วัดมาตรฐานและตรวจสอบงานด้านเทคโนโลยีสารสนเทศ และใช้มาตรฐาน ISO/IEC 27002 เพื่อแก้ไขปัญหาด้านความปลอดภัยของข้อมูลสารสนเทศ ลดความเสี่ยงที่จะเกิดขึ้นกับสารสนเทศภายในองค์กร [2] และการบริหารจัดการเหตุการณ์ไม่ปกติเป็นส่วนสำคัญที่ช่วยสนับสนุนการทำงานของระบบ [3] จึงมีการนำกระบวนการจัดการเหตุการณ์ไม่ปกติ และการจัดการปัญหาสำหรับการสนับสนุนบริการตามมาตรฐานไอทิล มาใช้ปรับปรุงกระบวนการจัดการบริการสารสนเทศขององค์กรให้มีประสิทธิภาพสูงสุด

งานวิจัยนี้จึงมุ่งเน้นพัฒนาระบบในการจัดการกับขั้นตอนกิจกรรมที่เกี่ยวข้องกับข้อตกลงระดับในการให้บริการ กระบวนการจัดการเหตุการณ์ไม่ปกติและปัญหาภายใต้มาตรฐานไอทิลกับระบบสนับสนุนการให้บริการอินเทอร์เน็ต เพื่อช่วยอำนวยความสะดวก ปรับปรุงคุณภาพและมาตรฐานการให้บริการ เพิ่มความพึงพอใจให้แก่ผู้ใช้บริการ และเพิ่มประสิทธิภาพในการปฏิบัติงาน รวมไปถึงการวิเคราะห์ข้อมูลเพื่อช่วยสนับสนุนการตัดสินใจสำหรับผู้บริหารในการหาแนวทางปรับปรุงคุณภาพการให้บริการให้เป็นไปได้ในทางที่ดีขึ้นในอนาคต

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 แนวคิดเกี่ยวกับไอทิล [4]

Information Technology Infrastructure Library (ITIL) หรือ ไอทิล เป็นชุดของหนังสือที่อธิบายวิธีปฏิบัติที่ดีที่สุด (Best Practices) ในการพัฒนาระบบบริหารจัดการงานให้บริการด้านระบบสารสนเทศที่ได้รับความนิยมอย่างกว้างขวาง มีประสิทธิภาพและประสิทธิผลชัดเจน โครงสร้างของไอทิลเวอร์ชัน 3 ประกอบด้วยมาตรฐานหลัก 5 ประการ แสดงดังภาพที่ 1



ภาพที่ 1: โครงสร้างของไอทิล เวอร์ชัน 3

2.1.1 กลยุทธ์งานบริการ (Service Strategy) เป็นการกำหนดแนวทางการบริหารการบริการ (Service Management) ในการกำหนดและบริหารนโยบายเชิงปฏิบัติ และกระบวนการในการบริหารการบริการอย่างครบวงจร เพื่อให้ทราบถึงโอกาสในการพัฒนาระบบการให้บริการแก่ตลาดธุรกิจให้สอดคล้องกับความต้องการของผู้ใช้บริการในองค์กรและลูกค้า เพื่อให้เกิดผลลัพธ์ของวิธีการบริการที่ดีที่สุด การออกแบบวิธีการนำระบบที่มีประสิทธิภาพไปใช้งาน

2.1.2 การออกแบบงานบริการ (Service Design) เป็นการออกแบบกิจกรรมในกระบวนการให้บริการ รวมทั้งการพัฒนากลยุทธ์และวิธีการบริหารจัดการระบบบริการ

2.1.3 การส่งมอบงานบริการ (Service Transition) เป็นการดำเนินการเพื่อให้ได้ผลลัพธ์ของการบริการที่ดีที่สุด รวมทั้งสร้างวิธีการบริการใหม่ ๆ ตลอดจนการปรับปรุงวิธีการบริการเดิม โดยมีข้อมูลเชื่อมโยงไปถึงการดำเนินงานบริการ (Service Operation)

2.1.4 การดำเนินงานบริการ (Service Operation) เป็นกิจกรรมที่จำเป็นต่อการปฏิบัติงาน เพื่อให้บรรลุผลสำเร็จในการดูแลรักษาหน้าที่การทำงานหรือบริการที่เป็นไปตามข้อตกลงตามพันธะสัญญาบริการ (Service Level Agreement) ที่มีต่อลูกค้า

2.1.5 การพัฒนางานด้านบริการ (Continual Service Improvement) เน้นขีดความสามารถที่จะทำให้เกิดการปรับปรุงการให้บริการที่มีคุณภาพอยู่แล้วให้มีความต่อเนื่องต่อความต้องการทางธุรกิจ

จุดแข็งของไอทิล คือ วิธีการอธิบายกระบวนการกับกิจกรรมที่แตกต่างกันได้และผังการทำงานที่ใช้ดำเนินงานเพื่อให้บรรลุเป้าหมายตามที่องค์กรวางไว้ ส่วน COBIT จะเน้นในเรื่องของการควบคุมและตรวจสอบการใช้สารสนเทศภายในองค์กรเพื่อให้คุ้มค่าต่อการลงทุน ส่วนมาตรฐาน ISO/IEC 27002 ใช้สำหรับการรักษาความปลอดภัยข้อมูล [1]

โดยงานวิจัยนี้เป็นการนำมาตรฐานไอทิลมาประยุกต์กับกระบวนการทำงานในการบริหารจัดการและแก้ไขเหตุการณ์ไม่ปกติในการให้บริการอินเทอร์เน็ต จึงนำข้อตกลงระดับการบริการ ในส่วนของการออกแบบงานบริการ และกระบวนการจัดการเหตุการณ์ไม่ปกติและจัดการปัญหา ในส่วนการดำเนินงานบริการ ซึ่งเป็นส่วนประกอบภายใต้มาตรฐานไอทิลมาประยุกต์ใช้กับกระบวนการทำงานของระบบสนับสนุนการให้บริการอินเทอร์เน็ต

2.2 งานวิจัยที่เกี่ยวข้อง

ขจรวุฒิ [3] นำเสนอกระบวนการจัดการเหตุการณ์ไม่ปกติและการจัดการปัญหาสำหรับการสนับสนุนบริการตามมาตรฐานไอทิล โดยเน้นพัฒนาระบบอำนวยความสะดวกสำหรับการปฏิบัติงาน โดยมีการออกแบบขั้นตอนการดำเนินงานของกระบวนการจัดการอินซิเด็นท์และการจัดการปัญหาในการให้บริการทางด้านเทคโนโลยีสารสนเทศ (IT) ร่วมกับเทคโนโลยีเว็บเซอร์วิสตามกระบวนการธุรกิจ ความยืดหยุ่นของระบบขึ้นอยู่กับความต้องการและประเภทการให้บริการไอทีของแต่ละองค์กร

นพพล [4] นำเสนอระบบการจัดการ โครงแบบสำหรับการสนับสนุนบริการตามมาตรฐานไอทิล โดยเน้นไปที่การจัดการ โครงแบบ เป็นกระบวนการจัดการ โครงแบบให้ข้อมูลของคอนฟิгурเรชันไอเท็มหรือซีไอมีความถูกต้อง เพื่อสนับสนุนกระบวนการจัดการบริการในการสนับสนุนบริการ มีการออกแบบและพัฒนาระบบจัดการ โครงการ โดยใช้เทคโนโลยีเว็บเซอร์วิส พบว่า ความสัมพันธ์ระหว่างซีไอกับเจ้าของซีไอ ในกระบวนการจัดการกับเหตุการณ์ไม่ปกติและกระบวนการจัดการปัญหา สามารถนำข้อมูลมาช่วยในการพิจารณาการส่งมอบปัญหาให้กับเจ้าหน้าที่เพื่อแก้ไขปัญหาที่เกิดขึ้นกับซีไอได้

ภาสิต [5] นำเสนอกระบวนการจัดการการเปลี่ยนแปลง สำหรับการสนับสนุนบริการตามมาตรฐานไอทิล โดยเน้นไปที่

การจัดการการเปลี่ยนแปลง เป็นการสร้างมาตรฐานในการทำการเปลี่ยนแปลงสถานะของไอซีต่าง ๆ ที่เกี่ยวข้องกับการให้บริการเทคโนโลยีสารสนเทศให้สอดคล้องกับความต้องการทางธุรกิจ การเปลี่ยนแปลงอาจเป็นผลมาจากกระบวนการจัดการปัญหา มีการออกแบบและพัฒนาระบบอำนวยความสะดวกให้องค์กรปฏิบัติงานของกระบวนการการเปลี่ยนแปลงตามมาตรฐานไอทิล โดยใช้เทคโนโลยีเว็บเซอร์วิส และการจำลองแบบกระบวนการธุรกิจ ซึ่งใช้แนวคิดกระแสน้ำ พบว่าสามารถเชื่อมโยงและให้บริการข้อมูลการเปลี่ยนแปลงของซีไอในระบบ รวมไปถึงข้อมูลข้อกำหนดการเปลี่ยนแปลงล่วงหน้า

Shaohua Z. และคณะ [6] นำเสนอการบูรณาการกระบวนการไอทิลภายในองค์กรบนพื้นฐานของเว็บเซอร์วิส โดยมีการวิเคราะห์การแก้ไขปัญหามาเสนอโครงสร้างซึ่งเป็นการบูรณาการร่วมกันของกระบวนการไอทิลและระบบในองค์กรกับเว็บเซอร์วิส เน้นการใช้ประโยชน์มาตรฐานในการทำงานร่วมกันของเว็บเซอร์วิสกับกระบวนการไอทิล ซึ่งสามารถอำนวยความสะดวกในการทำงานร่วมกันและการแบ่งปันองค์ความรู้ผ่านเครือข่ายอินเทอร์เน็ต พบว่า การใช้คุณสมบัติการทำงานร่วมกันของเว็บเซอร์วิส ในการรวมหลาย ๆ แอปพลิเคชันไว้ด้วยกัน ทำให้การนำสถาปัตยกรรมการบูรณาการกระบวนการของไอทิลไปใช้ได้ง่ายขึ้น

3. วิธีการดำเนินงาน

งานวิจัยนี้ได้ศึกษา ออกแบบ และพัฒนาระบบการจัดการที่เกี่ยวข้องกับข้อตกลงระดับการให้บริการ กระบวนการจัดการเหตุการณ์ไม่ปกติและปัญหาภายใต้มาตรฐานไอทิลกับระบบสนับสนุนการให้บริการอินเทอร์เน็ต ผู้วิจัยได้นำมาตรฐานไอทิลมาประยุกต์ใช้กับกระบวนการรับแจ้งเหตุเสียในการให้บริการอินเทอร์เน็ตซึ่งเป็นส่วนหนึ่งของระบบ นำมาสู่การวิเคราะห์ความต้องการของระบบ ดังนี้

3.1 Service Design

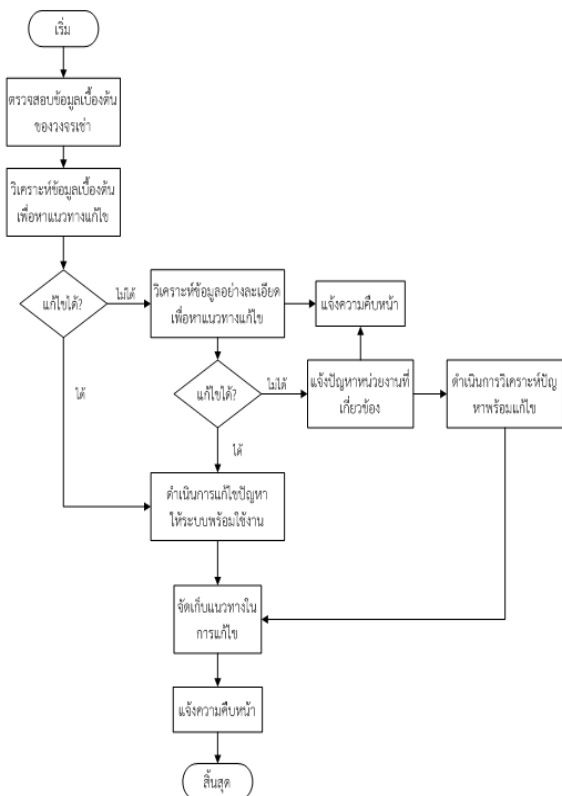
- การระบุระยะเวลาของการแก้ไขเหตุเสียภายในช่วงเวลาที่กำหนดของวงจรเช่าแต่ละวงจร หรือข้อตกลงการให้บริการ (SLA) ในการให้บริการตามสัญญาที่ตกลงไว้กับผู้ให้บริการ
- บันทึกและติดตามผลการปฏิบัติงานว่าเป็นไปตามข้อตกลงการให้บริการที่กำหนดหรือไม่

- จัดทำรายงานสรุปผลการแก้ไขเหตุเสียมตามข้อตกลงการให้บริการ

3.2 Service Operation

3.2.1 กระบวนการจัดการเหตุการณ์ไม่ปกติ (Incident Management) เป็นกระบวนการการแก้ไขเหตุการณ์ในการให้บริการที่ผิดปกติ ซึ่งเป็นสาเหตุทำให้การให้บริการหยุดชะงัก เน้นการกู้ระบบงานและความสามารถในการให้บริการให้กลับคืนสู่สภาวะปกติให้เร็วที่สุด เริ่มจากการรับแจ้งปัญหาหรือเหตุเสียมเกี่ยวกับเหตุขัดข้องในการใช้บริการอินเทอร์เน็ต จากนั้นทำการบันทึกรายละเอียดของปัญหา จัดประเภท และกำหนดระดับความสำคัญของเหตุการณ์ไม่ปกติที่เกิดขึ้น

จากนั้นทำการรวบรวมและวิเคราะห์ข้อมูลปัญหา เพื่อหาแนวทางแก้ไขเหตุการณ์ไม่ปกติที่เกิดขึ้น พร้อมรายงานความคืบหน้าของเหตุการณ์ไม่ปกติ สรุปผลข้อมูลของการแก้ไขปัญหา และบันทึกข้อมูลสถานะล่าสุดของเหตุการณ์ไม่ปกติ พร้อมทั้งปิดงานถ้าเหตุการณ์ไม่ปกตินั้นได้รับการแก้ไขให้สามารถให้บริการได้ตามปกติแล้ว ภาพที่ 2



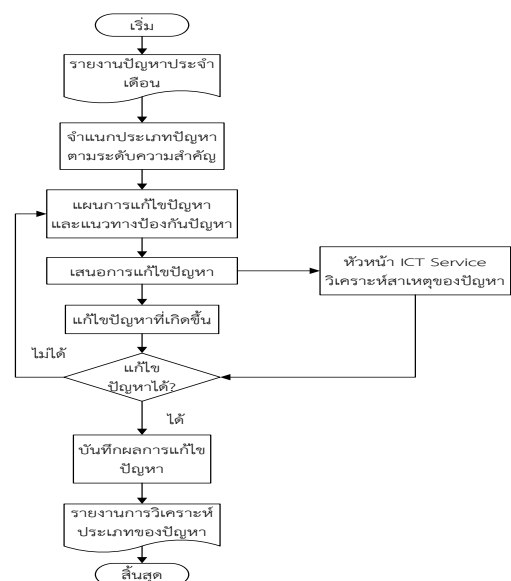
ภาพที่ 2: กระบวนการแก้ไขเหตุการณ์ไม่ปกติ

โดยได้มีการวิเคราะห์ห้ความต้องการของกระบวนการจัดการเหตุการณ์ไม่ปกติ ซึ่งมีรายละเอียดดังนี้

- การรับแจ้งและบันทึกข้อมูลเหตุการณ์ไม่ปกติ
- การจัดหมวดหมู่ของเหตุการณ์ไม่ปกติ โดยพิจารณาจากอาการเสียมของวงจรเช่า
- การจำแนกระดับความสำคัญของเหตุการณ์ไม่ปกติตามผลกระทบและความเร่งด่วนของเหตุการณ์ไม่ปกติในการแก้ไข
- การกู้คืนระบบ โดยการแก้ไขปัญหาของเหตุการณ์ไม่ปกติ เพื่อให้ระบบสามารถกลับมาทำงานได้ปกติภายในระยะเวลาตามข้อตกลงระดับบริการที่กำหนด
- การปิดงานเหตุการณ์ไม่ปกติ
- การสรุปรายงานและสถานะของเหตุการณ์ไม่ปกติ

3.2.2 ส่วนของกระบวนการจัดการปัญหา (Problem Management) เป็นการจัดการแก้ไขปัญหาจากสาเหตุที่แท้จริง และสามารถเตรียมความพร้อมในการป้องกันปัญหาก่อนที่จะเกิดขึ้นมาอีก เริ่มจากการนำรายงานบันทึกรายละเอียดเหตุการณ์ไม่ปกติที่เกิดขึ้นในแต่ละเดือนมาวิเคราะห์สาเหตุและจำแนกประเภทของปัญหาที่เกิดขึ้นบ่อย ๆ เพื่อจัดเป็นหมวดหมู่ จากนั้นทำการวิเคราะห์ ซึ่งอาจมีการเปลี่ยนแปลงระบบ เพื่อป้องกันไม่ให้อันนั้นเกิดขึ้นอีก แล้วทำการตรวจสอบว่าสามารถแก้ไขปัญหานั้นได้หรือไม่ ถ้าแก้ไขได้แล้วให้ทำการออกรายงาน แต่ถ้าแก้ไขไม่ได้ให้ทำการทบทวนใหม่ ดังภาพที่ 3

3



ภาพที่ 3: กระบวนการจัดการปัญหา

โดยได้มีการวิเคราะห์ความต้องการของกระบวนการจัดการปัญหา ซึ่งมีรายละเอียดดังนี้

- การนำข้อมูลจำแนกปัญหาประจำเดือนมาวิเคราะห์ เพื่อหาสาเหตุที่แท้จริงของปัญหา

- การสรุปการแก้ไขปัญหามาเป็นรายงาน

ผู้วิจัยได้นำความต้องการข้างต้นของระบบมาใช้ในการพัฒนาระบบ โดยระบบจะทำงานอยู่ในลักษณะของเว็บแอปพลิเคชัน เพื่อความสะดวกในการเข้าใช้งานและการเข้าถึงข้อมูลของผู้ใช้งาน และยังมีคามยืดหยุ่นในการพัฒนาปรับปรุงระบบให้เหมาะสมกับกระบวนการทำงานที่อาจมีการเปลี่ยนแปลงในอนาคต

4. ผลการดำเนินงาน

4.1 ผลการพัฒนาระบบ

จากแนวคิดและวิธีการดำเนินการวิจัยที่ได้นำเสนอ ผู้วิจัยได้พัฒนาระบบ โดยระบบมีความสามารถดังต่อไปนี้

- บันทึกข้อมูลเหตุการณ์ไม่ปกติของวงจรเช่าอินเทอร์เน็ต ตัวอย่างหน้าจอบันทึกเหตุการณ์ไม่ปกติ ดังภาพที่ 4

ภาพที่ 4: ตัวอย่างหน้าจอบันทึกเหตุการณ์ไม่ปกติ

- กำหนดระดับความสำคัญของเหตุการณ์ไม่ปกติ โดยพิจารณาตามผลกระทบและความเร่งด่วนของเหตุการณ์ไม่ปกติ และเป็นไปตามข้อตกลงระดับบริการ ดังภาพที่ 5

- สามารถแสดงเหตุการณ์ไม่ปกติที่ยังไม่ดำเนินการให้ระบบสามารถกลับมาใช้งานได้ปกติได้หรือเหตุการณ์กำลังอยู่ในขั้นตอนของการติดตามแก้ไขปัญหา

- จัดเก็บประวัติเหตุการณ์ไม่ปกติของแต่ละวงจรเช่า เพื่อนำไปใช้เป็นข้อมูลในการหาสาเหตุที่แท้จริงของปัญหา ในกรณีที่ยังเช่นนั้น ๆ มีเหตุการณ์ไม่ปกติเกิดขึ้นบ่อยครั้ง

- สามารถสืบค้นประวัติเหตุการณ์ไม่ปกติของแต่ละวงจรเช่าได้

- สามารถปิดงานของเหตุการณ์ไม่ปกติพร้อมระบุสถานะของเหตุการณ์ดังกล่าวได้ ระยะเวลาที่ใช้ในการแก้ไขเหตุสียกลับคืน และสถานะของข้อตกลงระดับบริการ (SLA) ว่าเป็นตามข้อตกลงระดับบริการหรือไม่

- รายงานสรุปพร้อมแสดงกราฟของเหตุการณ์ไม่ปกติในแต่ละเดือนว่าเป็นไปตามข้อตกลงระดับบริการหรือไม่ ดังภาพที่ 6

- รายงานสรุปพร้อมแสดงกราฟของเหตุการณ์ไม่ปกติในแต่ละเดือน โดยแยกตาม อาการเสีย สาเหตุของเหตุการณ์ไม่ปกติ พื้นที่ที่ติดตั้งวงจรเช่า ระดับความสำคัญ และสถานะของเหตุการณ์ไม่ปกติ

- รวบรวมเหตุการณ์ไม่ปกติประจำเดือนมาทำการวิเคราะห์ โดยใช้เกณฑ์ความถี่ของปัญหาที่เกิดขึ้นร่วมด้วย เพื่อหาสาเหตุที่แท้จริงของปัญหา พร้อมทั้งจัดเก็บ และหาวิธีป้องกันเพื่อไม่ให้ปัญหานั้นเกิดขึ้นอีก

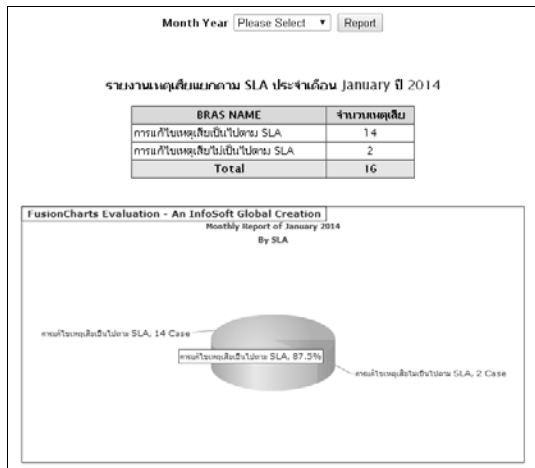
View All ADSL Fault							
Ticket No.	S/O	Account	อาการเสีย	วันเวลาที่รายงาน	Due By	Status	Priority
TF2014010027	A000001774	a000001774@superior	เข้า WEB ไม่ได้	2013-11-04 08:00:00	2013-11-05 08:00:00	Closed	Medium
TF2014010026	A000007465	a000007465@superior	Connect ไม่ได้	2013-11-05 09:29:00	2013-11-06 09:29:00	Closed	High
TF2014010025	A000002536	a000002536@superior	Connect ไม่ได้	2013-11-05 08:00:00	2013-11-06 08:00:00	Closed	High
TF2014010024	A000007673	a000007673@superior	เข้า WEB ไม่ได้	2013-11-11 08:00:00	2013-11-12 08:00:00	Closed	Medium
TF2014010023	A000001173	prt962284@regular	เข้า WEB ไม่ได้	2013-11-11 08:39:00	2013-11-15 08:39:00	Closed	Low
TF2014010022	A000007677	a000007677@superior	เข้า WEB ไม่ได้	2013-11-13 18:52:00	2013-11-14 18:52:00	Closed	Medium
TF2014010021	A000003933	a000003933@superior	Connect ไม่ได้	2013-11-14 11:54:00	2013-11-15 11:54:00	Closed	High
TF2014010020	A00000507	b000000001@basic	ใช้ Smart Phone ไม่ได้	2013-11-18 20:25:00	2013-11-19 20:25:00	Closed	Low
TF2014010019	A000003965	a000003965@superior	เข้า WEB ไม่ได้	2013-11-21 08:56:00	2013-11-22 08:56:00	Closed	Medium
TF2014010018	A000005454	a000005454@superior	Speed ไม่ได้	2013-11-20 09:13:00	2013-11-21 09:13:00	Closed	Medium
TF2014010017	A000007550	a000007550@superior	Speed ไม่ได้	2013-11-26 14:26:00	2013-11-26 22:26:00	Closed	High
TF2014010016	A000005362	a000005362@superior	เข้า WEB ไม่ได้	2013-11-01 08:00:00	2013-11-02 08:00:00	Closed	Medium
TF2014010015	A000008517	a000008517@superior	Speed ไม่ได้	2014-01-22 11:02:00	2014-01-23 11:02:00	Closed	Medium
TF2014010014	A000008565	a000008565@superior	เข้า WEB ไม่ได้	2014-01-20 20:20:00	2014-01-21 20:20:00	Closed	Medium

ภาพที่ 5: ตัวอย่างหน้าจอแสดงการกำหนดระดับความสำคัญ

4.2 ผลการประเมินประสิทธิภาพระบบ

เมื่อพัฒนาระบบเสร็จเรียบร้อยแล้ว จึงทำการประเมินประสิทธิภาพระบบ โดยผู้เชี่ยวชาญ 5 ด้าน คือด้านความสามารถในการทำงานได้ตรงตามความต้องการ ด้านหน้าที่ของระบบ ด้านการใช้งานระบบ ด้านประสิทธิภาพ และ

ด้านความปลอดภัยของระบบ ผลจากการประเมินประสิทธิภาพของระบบได้ค่าเฉลี่ยโดยรวมเท่ากับ 4.20 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.51 แสดงให้เห็นว่าระบบที่พัฒนาขึ้นมีประสิทธิภาพการใช้งานโดยรวมอยู่ในระดับดี สามารถนำไปใช้งานได้



ภาพที่ 6: ตัวอย่างหน้าจอแสดงรายงานแยกตามเงื่อนไข SLA

5. บทสรุป

งานวิจัยนี้ได้ออกแบบและพัฒนาระบบสนับสนุนการให้บริการอินเทอร์เน็ต โดยนำกระบวนการจัดการระดับบริการ การจัดการเหตุการณ์ไม่ปกติและการจัดการปัญหาตามมาตรฐานไอทิล และใช้เทคโนโลยีเว็บแอปพลิเคชันเป็นแนวทางในการพัฒนาระบบให้สอดคล้องกับกระบวนการธุรกิจการให้บริการอินเทอร์เน็ต การแก้ไขให้ระบบกลับมาพร้อมให้บริการอย่างรวดเร็วที่สุดหลังจากมีการรับแจ้งเหตุการณ์ไม่ปกติจากผู้ให้บริการภายใต้ข้อตกลงระดับบริการ มีการติดตามแจ้งความคืบหน้า รวบรวม และวิเคราะห์หาสาเหตุที่แท้จริง จะช่วยให้การให้บริการมีคุณภาพและประสิทธิภาพมากยิ่งขึ้น ลดความเสี่ยงที่จะทำให้เกิดความเสียหายต่อองค์กรผู้ให้บริการ ทั้งยังสามารถนำข้อมูลดังกล่าวไปใช้สนับสนุนการตัดสินใจของผู้บริหารได้อีกด้วย

เอกสารอ้างอิง

- [1] Shamsul S., Mohammad S. and Masarat A. Combining ITIL, "COBIT and ISO/IEC 27002 in Order to Design a Comprehensive IT Framework in Organizations," *Second Asia International Conference on Modeling & Simulation*, Kuala Lumpur Malaysia, 13-15 May 2008, 749-753.
- [2] Cusick, J.J.; Ma, G., "Creating an ITIL inspired Incident Management approach: Roots, response, and results". In *Procs. of the Network Operations and Manag. Symp. Workshops*, pp: 142-148, 2010.
- [3] ขจรวุฒิ น้อยอนุสนธิกุล, "กระบวนการจัดการอินซิเดนต์ และการจัดการปัญหาสำหรับการสนับสนุนบริการตามมาตรฐานไอทิล," *วิทยานิพนธ์ปริญญาโทบริหารธุรกิจ สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย*, 2551.
- [4] นพดล สิทธิเดชพร, "ระบบจัดการโครงสร้างสำหรับการสนับสนุนบริการตามมาตรฐานไอทิล," *วิทยานิพนธ์ปริญญาโทบริหารธุรกิจ สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย*, 2551.
- [5] ภาวิชิต บุญเกียรติ, "กระบวนการจัดการการเปลี่ยนแปลงสำหรับการสนับสนุนบริการตามมาตรฐานไอทิล," *วิทยานิพนธ์ปริญญาโทบริหารธุรกิจ สาขาวิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย*, 2551.
- [6] Shaohua Z., Peng W., Zhigang D. and Yuwei Z., "Organization ITIL Process Integration based on Web Services," *World Congress on Software Engineering*, Xiamen China, 19-21 May 2009, 412-416.

ระบบสารสนเทศเพื่อสนับสนุนการตัดสินใจ สำหรับผู้บริหารการประปานครหลวง

Decision Support System for Metropolitan Waterworks Authority Manager

อัยพรณ เอโกบล (Aiyapan Eagobon)¹ และ เนื่องวงศ์ ทวยเจริญ (Nuengwong Tuaycharoen)²

¹ ฝ่ายมาตรฐานวิศวกรรมและสารสนเทศภูมิศาสตร์ กองสารสนเทศภูมิศาสตร์ ส่วนสารสนเทศระบบท่อการประปานครหลวง

² ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยธุรกิจบัณฑิตย์

staykung@gmail.com, nuengwong.tun@dpu.ac.th

บทคัดย่อ

ในปัจจุบัน การสรุปข้อมูลประจำเดือนให้ผู้บริหารของการประปานครหลวงอยู่ในลักษณะของเอกสาร ซึ่งเวลาที่ใช้ในการรวบรวมข้อมูลเป็นไปอย่างล่าช้าและการแก้ไขเป็นไปอย่างยากลำบาก เนื่องจากข้อมูลมีจำนวนมากและมาจากหลายหน่วยงาน ผู้วิจัยจึงเห็นว่าการนำเอาเทคโนโลยีสารสนเทศภูมิศาสตร์และเทคโนโลยีเว็บเข้ามาใช้ เพื่อพัฒนาระบบให้บริการแผนที่จะทำให้เกิดประสิทธิภาพของการทำงานมากยิ่งขึ้น โดยวัตถุประสงค์ของงานวิจัยคือเน้นการแสดงผลข้อมูลต่างๆ ของทางภาครัฐและหน่วยงานของการประปานครหลวงเองที่สามารถจะนำมาเป็นประโยชน์ต่อการสนับสนุนการทำงานของการประปานครหลวง โดยใช้วิธีการนำข้อมูลที่มีการประปานครหลวงมีอยู่แล้วนำมาใช้ให้เกิดประโยชน์สูงสุด ผลของการดำเนินการวิจัยพบว่าเวลาในการเข้าถึงข้อมูลมีประสิทธิภาพทางเวลาเท่ากับ $O(\log_2 n)^2$ ดังนั้นระบบที่พัฒนาขึ้นสามารถทำงานได้อย่างมีประสิทธิภาพและเป็นประโยชน์ต่อการประปานครหลวงเป็นอย่างมาก

คำสำคัญ: ระบบสารสนเทศเพื่อสนับสนุนการตัดสินใจ เทคโนโลยีเว็บ เทคโนโลยีสารสนเทศภูมิศาสตร์ เทคโนโลยีสำหรับหน่วยงานราชการ

Abstract

Currently, Working on monthly reports for the managers of the Metropolitan Waterworks Authority (MWA) uses paper-based method. The officers require a long period of time and much effort to gather necessary information from many data sources, and the report is difficult to correct if needed. We propose Decision Support System for Metropolitan Waterworks Authority

Manager or MWA GIS-DSS, which is developed with the GIS technology along with various web technologies to facilitate the reporting process. The objective of the research focuses on the real-time data displaying from various sources on the map, including the data from government agencies and from the Metropolitan Water Authority. The system can be useful greatly to support the work of the MWA by using the information that MWA already has. The results of the study show that time effectiveness of the system is approximately equal to $O(\log n)^2$. Therefore, the application can be used effectively to a great extent.

Keyword: DSS, Web Technology, GIS, e-government.

1. บทนำ

การประปานครหลวง (กปน.) ทำการแปรรูปน้ำดิบและผลิตน้ำเพื่อการอุปโภคบริโภคสำหรับประชาชนในกรุงเทพฯ และปริมณฑล ซึ่งกระบวนการต่างๆ นั้นอาศัยข้อมูลและความสามารถในการติดตามผลที่รวดเร็ว ทันเหตุการณ์ ในการบริหารจัดการระบบขนส่งน้ำ เพื่อสนับสนุนการตัดสินใจของผู้ปฏิบัติงานรวมถึงผู้บริหาร ในอดีต กปน.จัดเก็บข้อมูลท่อประปา (Pipe Line)ในรูปแบบกระดาษไข เมื่อท่อจ่ายครบอายุการใช้งาน และมีความจำเป็นต้องขุดวางท่อทดแทน จึงต้องแก้ไขแบบก่อสร้างจริง (As-Built Drawings) โดยการขุดกระดาศบริเวณเส้นท่อจ่ายที่ครบอายุการใช้งานและทำการวางท่อทดแทน ซึ่งสร้างความลำบากให้แก่ผู้ปฏิบัติงานเป็นอย่างยิ่ง บ่อยครั้งที่เกิดเหตุการณ์ท่อแตกรั่ว น้ำซึมผ่านผิวถนน จำเป็นต้องทำการปิดน้ำเพื่อทำการซ่อมแซมท่อ ซึ่งส่งผลกระทบต่อผู้ใช้น้ำในเชิงพาณิชย์เป็นอย่างมาก ด้วยปัญหาที่

กล่าวข้างต้นจึงก่อเกิดแนวคิดในการพัฒนาระบบสารสนเทศภูมิศาสตร์เพื่อสนับสนุนการตัดสินใจด้านวิศวกรรมแบบรวมศูนย์ (Centralizes System) ที่สามารถใช้ทรัพยากรร่วมกันภายในหน่วยงานบนอินเทอร์เน็ตและอินเทอร์เน็ตได้ โดยมีระบบสารสนเทศเพื่อสนับสนุนการตัดสินใจ สำหรับผู้บริหาร กปน. หรือ MWA GIS-DSS ที่ผู้วิจัยได้พัฒนาขึ้นเป็นส่วนประกอบหนึ่ง ที่สามารถแสดงข้อมูลเพื่อสนับสนุนการตัดสินใจสำหรับผู้บริหารทุกระดับ โดยแสดงข้อมูลแบบเวลาจริง (Real-Time) การพัฒนาระบบนี้จะสามารถลดค่าใช้จ่ายจากการลงทุนพัฒนาระบบที่ซ้ำซ้อนกันได้

ในด้านการแสดงผลข้อมูล ระบบ MWA GIS-DSS จะแสดงผลข้อมูลต่าง ๆ บนแผนที่ ซึ่งเป็นข้อมูลเพื่อการบริหารจัดการทางด้านวิศวกรรมและการก่อสร้างระบบประปา เพื่อให้ผู้บริหารตั้งแต่ระดับต้นถึงระดับสูงสามารถมองเห็นภาพรวมผ่านชั้นข้อมูลต่างๆ จนนำไปสู่การตัดสินใจและผลักดันนโยบายการใช้งานได้ ระบบนี้ยังได้นำไปสู่การสนับสนุนการลงทุนด้านระบบสารสนเทศภูมิศาสตร์ รวมถึงสร้างเครื่องมือต้นแบบในการแลกเปลี่ยนข้อมูลให้กับระบบงานต่างๆ เช่น CIS, WLMA, SAP และ CRM ให้ได้รับข้อมูลที่ทันสมัยต่อไปในอนาคต

2. ทฤษฎีที่และงานวิจัยและเกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้อง

ระบบที่พัฒนาขึ้น ได้ใช้เทคโนโลยีต่าง ๆ ดังต่อไปนี้

2.1.1 SOAP (Simple Object Access Protocol)

SOAP เป็นโปรโตคอลที่ใช้ในการติดต่อสื่อสารหรือแลกเปลี่ยนข้อมูลระหว่าง web services โดยติดต่อผ่านทาง HTTP โปรโตคอลซึ่งอาศัยรูปแบบของภาษา XML (Extensible Markup Language) ทำให้ Web services สามารถสื่อสารกันได้แม้ว่าจะอยู่บนเครื่องคอมพิวเตอร์คนละแพลตฟอร์มก็ตาม โดย SOAP ยังสามารถทำงานร่วมกับโปรโตคอลได้หลายชนิด เช่น HTTP, SMTP, FTP [1]

2.1.2 OGC (Open Geospatial Consortium)

OGC เป็นมาตรฐานที่ใช้ในการติดต่อสื่อสารหรือแลกเปลี่ยนข้อมูลสารสนเทศภูมิศาสตร์โดย OGC เกิดจากการรวมตัวกันของกลุ่มหน่วยงานภาครัฐและเอกชนและกลุ่มของ

บริษัทต่างๆในภาคอุตสาหกรรมโดย OGC ได้กำหนดมาตรฐานกลางที่ใช้ในการติดต่อสื่อสารกันโดยในปัจจุบันที่นิยมใช้กัน เช่น KML (Keyhole Markup Language), WMS (Web Map Service), WFS (Web Feature Service), WTS (Web Terrain Service) [2]

2.1.3 MVC (Model-View-Controller)

MVC เป็นสถาปัตยกรรมซอฟต์แวร์ซึ่งแบ่งหน้าที่การทำงานออกเป็น 3 ส่วนหลักๆ ได้แก่ 1.) M (Model) ทำหน้าที่ติดต่อกับฐานข้อมูลโดย Model รับคำสั่งมาจาก Controller ทั้งการ Update, Insert หรือ Delete เป็นต้น 2.) V (View) ทำหน้าที่แสดงผลส่วนที่ติดต่อกับผู้ใช้งานและสามารถนำข้อมูลในส่วน ของ Model ส่งมาแสดงผลได้โดยผ่าน Controller 3.) C (Controller) ทำหน้าที่คำนวณและประมวลผลต่างๆ ของระบบทั้งหมด [3]

2.1.4 ระบบสารสนเทศภูมิศาสตร์

ระบบสารสนเทศภูมิศาสตร์หรือเรียกอีกชื่อหนึ่งว่า GIS (Geographic Information System) คือระบบการทำงานเกี่ยวกับข้อมูลทางด้านแผนที่ที่สามารถนำข้อมูลเชิงพื้นที่มาวิเคราะห์ได้โดยใช้ข้อมูลเชิงพื้นที่ (Spatial Data) แบ่งออกได้เป็น 2 ประเภท คือ 1.) ข้อมูลเวกเตอร์ (Vector Data) ข้อมูลทางเวกเตอร์จะประกอบไปด้วย ข้อมูลจุด (Points) ใช้อ้างอิงตำแหน่งต่างๆ ในแผนที่เช่น วัด โรงเรียน บ้าน ข้อมูลเส้น (Lines) เป็นข้อมูลของจุดหลายๆ จุดต่อกัน เช่น แม่น้ำ คลอง ถนน 2.) ข้อมูลราสเตอร์ (Raster Data) จะเก็บข้อมูลเป็นรูปภาพชิ้นเล็กๆ มาเรียงต่อกันโดยอ้างอิงพิกัดจากตำแหน่งภูมิศาสตร์ เช่น ภาพถ่ายทางอากาศ (Aerial) ภาพถ่ายดาวเทียม (Satellite)[4]

2.1.5 AJAX (Asynchronous JavaScript and XML)

AJAX เป็นการทำงานในแบบ Asynchronous คือเป็นการโหลดส่วนของข้อมูลต่างๆที่ต้องการ โดยที่หน้าเว็บไซต์ไม่ต้องโหลดทั้งหมด ซึ่ง AJAX จะทำให้โปรแกรมประยุกต์มีประสิทธิภาพมากยิ่งขึ้นโดย AJAX จะทำการร้องขอข้อมูลที่ต้องการไปยังเซิร์ฟเวอร์และเมื่อเซิร์ฟเวอร์ตอบกลับมาก็จะให้ข้อมูลที่เรากำลังต้องการกลับมาด้วย โดย AJAX สามารถดึงข้อมูลในรูปแบบของ XML หรือ JSON ได้อีกด้วย [5]

2.1.6 JSON (Java Script Object Notation)

JSON เป็นรูปแบบของข้อมูลที่ใช้สำหรับการแลกเปลี่ยนข้อมูลผ่านเซิร์ฟเวอร์ ซึ่งมนุษย์สามารถเข้าใจได้ง่ายกว่าภาษา XML โดย JSON ถูกสร้างขึ้นด้วยภาษา Java Script จะใช้ [] แทน Array และใช้ {} แทน Hash แต่ละสมาชิกคั่นด้วย comma (,) และแต่ละ ชื่อสมาชิกคั่นด้วย colon (:) ซึ่งในปัจจุบัน JSON ถูกนิยมใช้งานร่วมกับ AJAX [6]

2.2 งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่มีการนำระบบสารสนเทศทางภูมิศาสตร์มาใช้ในประเทศไทยที่เกี่ยวกับหน่วยงานราชการ มีดังนี้

2.2.1 ระบบสนับสนุนการตัดสินใจเพื่อการพัฒนาการท่องเที่ยวหมู่เกาะสีชัง จังหวัดชลบุรี

โปรแกรมประยุกต์นี้พัฒนาขึ้นด้วย ASP.NET โดยใช้ภาษา VB ในการพัฒนาโปรแกรมประยุกต์และใช้โปรแกรม Microsoft Access เพื่อเก็บข้อมูลเชิงพื้นที่ (Spatial data) และข้อมูลตามลักษณะ (Attribute data) และใช้โปรแกรม ArcView ในการทำแผนที่และการวิเคราะห์ข้อมูลต่างๆ โดยงานวิจัยนี้พัฒนาขึ้นเพื่อให้นักท่องเที่ยวสามารถเรียกค้นข้อมูลในหมู่เกาะสีชังจังหวัดชลบุรี เพื่อใช้ประกอบการสนับสนุนในการวางแผนการท่องเที่ยวและสามารถกำหนดช่วงเวลากิจกรรมและความเหมาะสมกับช่วงเวลาได้ โปรแกรมประยุกต์ขั้นนี้ยังสามารถสำรวจและรวบรวมข้อมูลพื้นฐานด้านทรัพยากรที่เกี่ยวข้องกับการท่องเที่ยว เช่น แหล่งท่องเที่ยว ชุมชนท้องถิ่น ประชากร เขตการปกครองโครงข่ายได้ [7]

2.2.2 การประยุกต์ระบบสารสนเทศภูมิศาสตร์สำหรับการวางแผนขยายสายส่ง

โครงการนี้ทำขึ้นเพื่อวางแผนวิธีการวิเคราะห์สายส่งไฟฟ้าของการไฟฟ้าส่วนภูมิภาค โดยนำเข้ามาพร้อมกับระบบสารสนเทศภูมิศาสตร์ของการไฟฟ้าส่วนภูมิภาคซึ่งเป็นระบบจัดเก็บ ประมวลผล วิเคราะห์ข้อมูลและสืบค้นข้อมูลของอุปกรณ์ไฟฟ้าทั้งหมดในเชิงของแผนที่โดยโปรแกรมนี้มีวัตถุประสงค์เพื่อลดต้นทุนในการก่อสร้างสายส่งของไฟฟ้าค่าบำรุงรักษาโดยโปรแกรมประยุกต์นี้สามารถวิเคราะห์โหลดโพล์ที่เรียกว่า LFAP (Load Flow Application Program Interface) ได้ถูกพัฒนาขึ้นเพื่อใช้วิเคราะห์แรงดันที่บัส กระแสในสายส่งและกำลังไฟฟ้าสูญเสียซึ่งทำให้ผู้ใช้สามารถวิเคราะห์

และวางแผนทางเลือกในการขยายสายส่งได้บนโปรแกรมประยุกต์โดยตรงทำให้ได้คำตอบที่พอเหมาะและเป็นไปได้ในทางปฏิบัติ [8]

3. รายละเอียดการพัฒนา

3.1 ภาพรวมของระบบสารสนเทศภูมิศาสตร์ของ กปน.

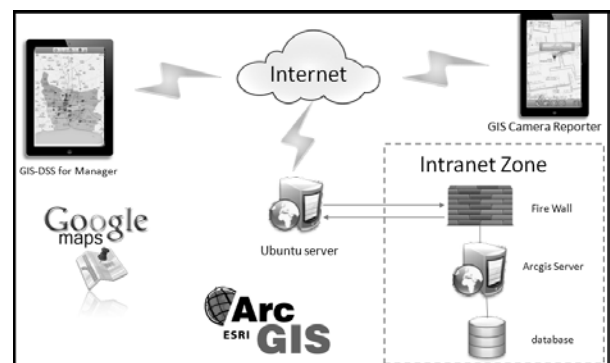
และ MWA GIS-DSS

กปน. ได้พัฒนาระบบสารสนเทศภูมิศาสตร์อย่างต่อเนื่องซึ่งประกอบด้วยระบบงานย่อยๆ แสดงในภาพที่ 1 ได้แก่

1.) GIS Camera Reporter ทำหน้าที่รับข้อมูลจากภาคสนามผ่านทางอินเทอร์เน็ต ซึ่งรองรับทั้งระบบปฏิบัติการ iOS และ Android โปรแกรมนี้ถูกพัฒนาขึ้นเพื่อใช้ภายใน กปน. ก่อนหน้านี้ โดยผู้ใช้สามารถดาวน์โหลดได้จาก App Store หรือ Google Play Store

2.) ArcGIS ทำหน้าที่ปรับปรุงข้อมูลบนแผนที่ฐานและชั้นข้อมูลด้านวิศวกรรมและการผลิตน้ำให้เกิดความทันสมัย และ

3.) MWA GIS-DSS ทำหน้าที่เชื่อมโยงฐานข้อมูลต่างๆ ภายใน กปน. รวมถึงฐานข้อมูลจากหน่วยงานต่างๆ ของรัฐ และแสดงชั้นข้อมูลบนแผนที่ฐาน รวมถึงทำการประมวลผลข้อมูลเชิงสถิติเพื่อสนับสนุนการตัดสินใจสำหรับผู้บริหาร และแสดงผลในรูปแบบกราฟ ซึ่งขอบเขตของบทความนี้ครอบคลุมเพียงระบบ MWA GIS-DSS เท่านั้น



ภาพที่ 1: โครงสร้างระบบสารสนเทศภูมิศาสตร์ กปน.

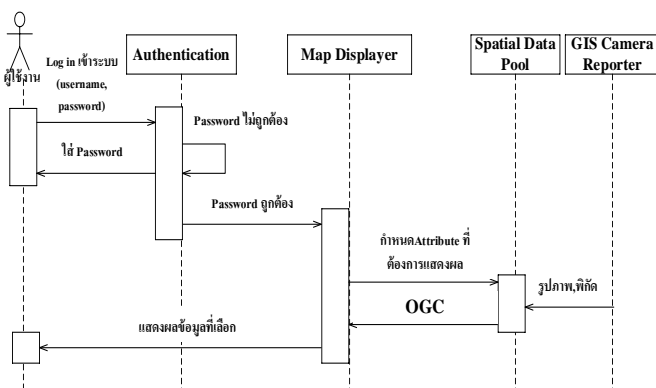
จากภาพที่ 1 การทำงานของระบบ MWA GIS-DSS เริ่มจากพนักงานของ กปน. รายงานสถานการณ์ต่างๆ จากภาคสนามเข้ามายังระบบโดยผ่าน Smartphone ที่ติดตั้งโปรแกรม GIS Camera Reporter เมื่อผู้ใช้งานส่งภาคสนามถ่ายรูปและใส่

รายละเอียดเหตุการณ์ เช่น เกิดเหตุท่อประปาแตกแล้วเพื่อที่จะรายงานสถานการณ์ต่างๆมายังระบบ โดยทางฝั่งของ GIS Camera Reporter ส่งข้อมูลมายัง Ubuntu Server ซึ่งข้อมูลประกอบไปด้วย ค่าพิกัด (Latitude, Longitude) ตำแหน่งจุดที่อยู่ปัจจุบันโดยอิงข้อมูลจาก GPS, รูปถ่าย และรายละเอียดอื่น ๆ

จากนั้น MWA GIS-DSS ซึ่งเป็นโปรแกรมสำหรับผู้บริหารของ กปน. ก็จะแสดงข้อมูลเป็นจุดที่เกิดเหตุ (Point) ของทางฝั่งของภาคสนามที่รายงานสถานการณ์เข้ามาแบบเวลาจริง นอกจากนี้ยังสามารถเลือกดูข้อมูลต่างๆ ที่ทางฝั่งของภาคสนามส่งเข้ามาได้อีกด้วย ซึ่งข้อมูลตรงนี้เหมาะสำหรับผู้บริหารเพื่อใช้สำหรับการตัดสินใจกับเหตุการณ์ต่างๆที่เกิดขึ้นได้ นอกจากนี้ ระบบ MWA GIS-DSS ยังสามารถแสดงชั้นข้อมูลต่างๆ ที่ได้เชื่อมโยงกับทางหน่วยงานภาครัฐต่างๆ ที่เป็นประโยชน์ต่อหน่วยงานของ กปน. อีกมากมาย และนำค่าทางสถิติมาสร้างกราฟ เพื่อนำเสนอข้อมูลที่เหมาะสมต่อการตัดสินใจสำหรับผู้บริหาร กปน. ในระดับต่างๆ

3.2 การออกแบบและพัฒนาโปรแกรม

3.2.1 แผนภาพซีเควนซ์ของระบบ



ภาพที่ 2: แผนภาพซีเควนซ์ของระบบ MWA GIS-DSS

ภาพที่ 2 แสดงแผนภาพซีเควนซ์เพื่อแสดงการทำงานของคลาสต่างๆ ในระบบ เมื่อผู้ใช้กรอกชื่อผู้ใช้และรหัสผ่านเข้าสู่ระบบผ่านระบบเว็บ MWA GIS-DSS แล้ว ระบบจะแสดงผลเป็นแผนที่ขึ้นมาให้เลือกชั้นข้อมูลที่ต้องการให้แสดง โดยระบบจะแบ่งข้อมูลออกเป็น 5 ส่วน คือ

1.)ข้อมูลทางด้านงานด้านวิศวกรรมโดยจะเน้นเฉพาะข้อมูลที่เกี่ยวข้องกับทางด้านวิศวกรรมของ กปน. และหน่วยงานภาครัฐที่เกี่ยวข้อง

2.)ข้อมูลทางด้านแหล่งน้ำ จะมีชั้นข้อมูลแหล่งน้ำต่างๆ ของ กปน. และชั้นข้อมูลของหน่วยงานภาครัฐต่างๆที่เกี่ยวข้อง

3.)ข้อมูลแจ้งเหตุการณ์ต่างๆแบบเวลาจริงโดยข้อมูลนี้ได้จากโปรแกรม GIS Camera Reporter

4.)ข้อมูลทางด้านแผนที่ของ กปน. ซึ่งจะมีข้อมูลอยู่ 3 จังหวัด ได้แก่ กรุงเทพฯ, นนทบุรี และสมุทรปราการ ข้อมูลนี้สามารถนำไปใช้ประโยชน์ในงานก่อสร้างของ กปน. เพื่อที่จะดูตำแหน่งแนวท่อ

5.)กราฟรายงานข้อมูลท่อประปาที่ กปน. มีอยู่ ซึ่งผู้ใช้สามารถมีปฏิสัมพันธ์กับกราฟได้

3.2.2 ซอฟต์แวร์ที่ใช้พัฒนา

ระบบของโปรแกรมประยุกต์นี้พัฒนาด้วยโปรแกรม Eclipse-PHP และเขียนด้วยภาษา PHP, HTML, JAVASCRIPT โดยมีฐานข้อมูลเป็น MySQL และใช้ XAMPP เป็นโปรแกรมเว็บเซิร์ฟเวอร์ นอกจากนี้ยังใช้เทคโนโลยีเว็บอีกมากมาย ได้แก่ JQuery, JQuery Mobile, Google Map API, HighCharts API และ Google Map utility-library V3 ในการพัฒนา

4. ผลการทดสอบ

4.1 สภาพแวดล้อมในการทดสอบ

ผู้วิจัยได้ทดสอบระบบบนเครื่องคอมพิวเตอร์ที่ติดตั้งระบบปฏิบัติการ Microsoft Windows 8.1 Professional โดยทดสอบที่บราวเซอร์ Google Chrome เพราะขอบเขตของงานวิจัยนี้เน้นที่ส่วน MWA GIS-DSS ที่มีหน้าที่รับข้อมูลจากโปรแกรมภาคสนามมาแสดงผลเท่านั้น

ช่วงเวลาของการทดสอบ เป็นปัจจัยหนึ่งต่อผลการทดสอบ เพราะช่วงเวลาในการทดสอบทำให้มีความแตกต่างในด้านจำนวนผู้ใช้งานในขณะนั้น และความเร็วของอินเทอร์เน็ต โดยในการทดสอบนี้ ผู้วิจัยได้ทำการทดสอบที่เวลาประมาณ 15.00 น. และใช้ Web Console ของ Fire Fox เป็นตัวช่วยในการจับเวลาในการทดสอบ

4.2 การทดสอบระบบตามหน้าที่ของระบบ

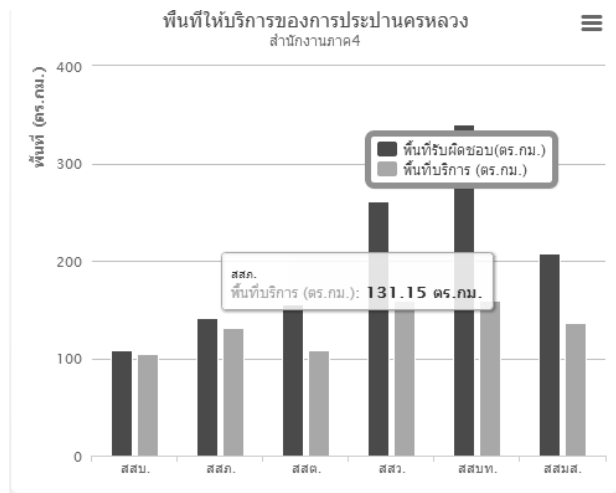
จากภาพที่ 3 แสดงหน้าจอหลักของระบบ MWA GIS-DSS ที่สามารถแสดงชั้นข้อมูล RTU ซึ่งจะแสดงค่าแรงดันน้ำของพื้นที่ที่ประปาให้บริการ โดยผู้บริหารระดับต้นและระดับกลางสามารถใส่ข้อมูลดังกล่าวเพื่อวางแผนหรือประเมินปัญหาที่อาจ

เกิดขึ้นในแต่ละจุดในอนาคต ดังแสดงในภาพที่ 3 และผู้บริหารในระดับสูงสามารถดูข้อมูลภาพรวมในลักษณะของกราฟดังภาพที่ 4 เพื่อใช้ประกอบการตัดสินใจ

ระบบจะประกอบด้วย 6 เมนูหลักแสดงเป็นไอคอนด้านบนได้แก่ 1.) เมนูแสดงตำแหน่งที่อยู่ตำแหน่งปัจจุบันบนแผนที่ 2.) เมนูค้นหาข้อมูล เช่น ถนน หมู่บ้าน จังหวัด บ้านเลขที่ โดยอิงข้อมูลจาก Google 3.) เมนูแสดงผลด้านวิศวกรรมหลัก 4.) เมนูข้อมูลด้านแหล่งน้ำต่างๆ 5.) เมนูเปลี่ยนแผนที่ฐาน 6.) เมนูแสดงข้อมูลในลักษณะของกราฟรายงาน ซึ่งแสดงตัวอย่างกราฟรายงานในภาพที่ 4 กราฟแสดงพื้นที่ให้บริการของ กปน.



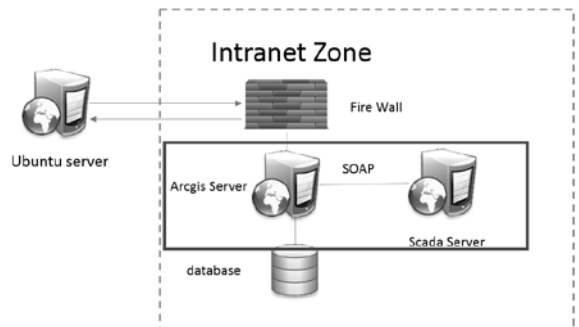
ภาพที่ 3: ค่าแรงดันบนพื้นที่ให้บริการของ กปน.



ภาพที่ 4: กราฟสำหรับผู้บริหารระดับสูงเพื่อเปรียบเทียบพื้นที่รับผิดชอบและพื้นที่ให้บริการแยกตามสาขา

โดยปกติแล้วช่องทางการสื่อสารข้อมูลของระบบภายในองค์กรของ กปน. จะสามารถติดต่อสื่อสารกันได้แค่ช่องทางเดียวเท่านั้น คือ ผ่านเครื่อง Ubuntu Server ซึ่งระบบสามารถดึง

ข้อมูลโดยใช้หลักการของ SOAP ที่พัฒนาด้วยภาษา Java Script ดังแสดงในภาพที่ 5



ภาพที่ 5: สถาปัตยกรรมการเชื่อมต่อแบบ SOAP [1] ของระบบ

ข้อมูลที่ได้จากการเชื่อมต่อข้อมูลกันแบบ SOAP ดังแสดงการเชื่อมต่อในภาพที่ 5 จะอยู่ในลักษณะของ JSON ดังแสดงได้ดังตัวอย่างรหัสในภาพที่ 6

```
[{"rtu_date": "2014-01-21 19:19:48", "rtu_name": "P19", "rtu_flow": "0", "rtu_pressure": "7", "rtu_lat": "13.708479093117", "rtu_lon": "100.342341143843"}, {"rtu_date": "2014-01-21 19:19:48", "rtu_name": "P21", "rtu_flow": null, "rtu_pressure": null, "rtu_lat": "13.577225360837", "rtu_lon": "100.570995150442"}, {"rtu_date": "2014-01-21 19:19:48", "rtu_name": "P32", "rtu_flow": "0", "rtu_pressure": "8.45", "rtu_lat": "13.66701", "rtu_lon": "100.52392"}]
```

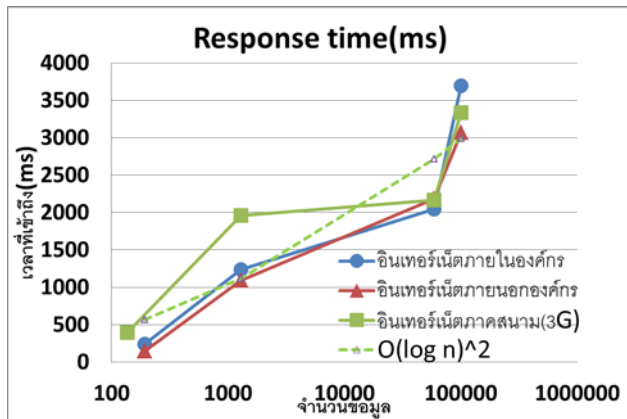
ภาพที่ 6: ตัวอย่างข้อมูล JSON ของระบบ

4.3 ประสิทธิภาพโดยรวมของระบบ

ภาพที่ 7 แสดงค่าเฉลี่ยของเวลาที่เว็บ MWA GIS-DSS ใช้ตั้งแต่การร้องขอจนกระทั่งแสดงผลครบถ้วนบนแผนที่ โดยแกน x แสดงถึงจำนวนข้อมูลที่แสดงบนแผนที่แบบ log scale และแกน y แสดงถึงเวลาที่ใช้ในการเข้าถึงและแสดงผลข้อมูลบนแผนที่ในหน่วย millisecond โดยข้อมูลในแต่ละจุดข้อมูลเป็นค่าเฉลี่ยจากการทดสอบร้องขอข้อมูลของ กปน. จำนวน 4 ครั้งจากการเชื่อมต่ออินเทอร์เน็ต 3 แบบคือ ภายในองค์กร ภายนอกองค์กร และอินเทอร์เน็ตภาคสนามซึ่งคือ 3G

จากภาพที่ 7 พบว่าระบบที่พัฒนาขึ้นสามารถแสดงข้อมูลจำนวน 1 แสนข้อมูลได้ภายในเวลาเพียง 4 วินาที ซึ่งเป็นเวลาที่ผู้ใช้เว็บทั่วไปปรับได้ [9] นอกจากนี้ เมื่อใช้วิธี Least Square ในการทำ Curve Fitting พบว่าระบบมีประสิทธิภาพทางเวลาเป็น

$O(\log_2 n)^2$ ซึ่งสามารถเทียบเคียงกราฟประสิทธิภาพดังแสดงในภาพที่ 7



ภาพที่ 7: กราฟแสดงค่าเฉลี่ยเวลาตั้งแต่ร้องขอจนถึงการแสดงผลบนเว็บ

4.4 การประเมินผลประสิทธิภาพของระบบจากผู้บริหาร

ผู้วิจัยได้ทำเก็บข้อมูลการประเมินผลความพึงพอใจด้านประสิทธิภาพของระบบจากผู้บริหารของ กปน. จำนวน 5 คน ได้คะแนนเฉลี่ย 4.20 จาก 5 คะแนน (5=พอใจมาก) ดังนั้นความพึงพอใจประสิทธิภาพโดยรวมของระบบอยู่ในระดับดี

5. สรุป

ระบบ MWA GIS-DSS ที่พัฒนาขึ้นนี้สามารถแสดงข้อมูลต่างๆ ที่เป็นประโยชน์เพื่อสนับสนุนการตัดสินใจในการทำงานด้านวิศวกรรมของ กปน. และยังสามารถรายงานสถานการณ์ต่างๆ ของผู้ที่ปฏิบัติงานภาคสนามโดยส่งข้อมูลมาแบบเวลาจริง โดยระบบมีประสิทธิภาพทางเวลาประมาณ $O(\log n)^2$ และการแสดงผลข้อมูลจำนวนหนึ่งแสนข้อมูลสามารถทำได้ภายในเวลาไม่ถึง 4 วินาที

ดังนั้น ระบบสนับสนุนการตัดสินใจของผู้บริหาร กปน. สามารถทำงานได้อย่างมีประสิทธิภาพและเป็นประโยชน์ต่อการทำงานและการตัดสินใจของผู้บริหารระดับต่าง ๆ ของ กปน.

6. กิตติกรรมประกาศ

งานวิจัยฉบับนี้สำเร็จได้ด้วยดี ด้วยความอนุเคราะห์อย่างสูงจากนายภาคภูมิ พิระชัย ที่ให้ข้อเสนอแนะเพิ่มเติมทำให้งานวิจัยฉบับนี้มีความถูกต้องสมบูรณ์ยิ่งขึ้น

ขอขอบคุณ นายอดิเทพ บาลารี่ ส่วนพัฒนาระบบสารสนเทศด้านวิศวกรรม กองพัฒนาระบบงานผลิตและวิศวกรรม กปน. เป็นผู้พัฒนาโปรแกรมประยุกต์บนมือถือ (GIS Camera Reporter)

เอกสารอ้างอิง

- [1] Christopher Kohlhoff. Evaluating SOAP for High Performance Business Applications: Real-Time Trading Systems [internet].2007 [cited 2014 Jan 15]. Available from:<http://www2003.org/cdrom/papers/alternate/P872/p872-kohlhoff.html>
- [2] <http://www.opengeospatial.org/>. [cited 2014 Jan 13].
- [3] Steve Burbeck. Applications Programming in Smalltalk-80(TM):How to use Model-View-Controller (MVC) [internet].[cited 2014 Jan 13]. Available from: <http://st-www.cs.illinois.edu/users/smarch/st-docs/mvc.html>
- [4] Kenneth E. Geographic Information Systems as an Integrating Technology: Context, Concepts, and Definitions [internet].[cited 2014 Jan 13]. Available from:<http://www.colorado.edu/geography/gcraft/notes/intro/intro.html>
- [5] Jesse James Garrett. Ajax: A New Approach to Web Applications [internet].[cited 2014 Jan 20]. Available from:<https://web.archive.org/web/20080702075113/http://www.adaptivepath.com/ideas/essays/archives/000385.php>
- [6] <http://allhandweb.com/?page=Blog.ShowBlogDetail&blogID=13/> (สืบค้นเมื่อวันที่ 22 มกราคม 2557).
- [7] โสภิต สร้อยสอดศรี, เสดิมศักดิ์ จารยะพันธุ์ และ ดุษฎี ชาญลิขิต. "ระบบสนับสนุนการตัดสินใจเพื่อการพัฒนาการท่องเที่ยวหมู่เกาะสี่จังหวัดชลบุรี". วิทยานิพนธ์ (วท.ม.) จุฬาลงกรณ์มหาวิทยาลัย 2549.
- [8] รัตติยา ยุทธวิชยานนท์, โสคติพงษ์ พิชัยสวัสดิ์ และ บัณฑิต เอื้ออาภรณ์. "การประยุกต์ระบบสารสนเทศภูมิศาสตร์สำหรับการวางแผนขยายสายส่ง". วิทยานิพนธ์ (วศ.ม.) จุฬาลงกรณ์มหาวิทยาลัย 2549.
- [9] Kissmetrics, "How loading time affects your bottom line". <http://goo.gl/kosva>

โครงร่าง Zachman สำหรับระบบการจัดการวิทยานิพนธ์หลักสูตรบัณฑิตศึกษาในสาขา ฟิสิกส์ มหาวิทยาลัยมหิดล

The Zachman Framework for Graduate Thesis and Dissertation Management System in Physics Program at Mahidol University

สันติ ปรีชาวิทย์ (Sunti Preechawit)¹ และภูมินท์ กิระวานิช (Phumin Kirawanich)²

^{1,2}สถาบันพัฒนาระบบการจัดการองค์กร ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ มหาวิทยาลัยมหิดล

¹ภาควิชาฟิสิกส์ คณะวิทยาศาสตร์ มหาวิทยาลัยมหิดล

¹ti_sunti@hotmail.com, ²phumin.kir@mahidol.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อสร้างโมเดลช่วยลดความเสี่ยงและความล้มเหลวในการพัฒนาระบบการจัดการวิทยานิพนธ์หลักสูตรบัณฑิตศึกษาในสาขาฟิสิกส์ มหาวิทยาลัยมหิดล โดยประยุกต์ใช้ความรู้ทางด้านสถาปัตยกรรมการจัดการองค์กร (Enterprise Architecture) ซึ่งโครงร่างที่นำมาใช้คือ โครงร่างแซคแมน (Zachman Framework) ร่วมกับภาษายูเอ็มแอล (UML) และพัฒนาระบบนี้ด้วยภาษาพีเอชพี (PHP) ร่วมกับฐานข้อมูลมายเอสคิวแอล (MySQL) ผลจากการพัฒนาพบว่าสามารถอธิบายรายละเอียดภาพรวมของการทำวิทยานิพนธ์ให้กับผู้ใช้ได้อย่างมีประสิทธิภาพ การประเมินความพึงพอใจระบบการจัดการวิทยานิพนธ์จากผู้ใช้จำนวน 20 คน ได้ผลประเมินค่าความพึงพอใจอยู่ในระดับดี

คำสำคัญ: สถาปัตยกรรมการจัดการองค์กร โครงร่างแซคแมน ระบบการจัดการวิทยานิพนธ์หลักสูตรบัณฑิตศึกษา

Abstract

The objective of this work is to create a model that helps reducing risks and failures in developing graduate thesis and dissertation management (GTDM) system in Physics Program at Mahidol University. Based on the comprehensive knowledge of Enterprise Architecture, the developed system uses the Zachman Framework incorporating with Unified Modeling Language (UML), Personal Home Page (PHP), and MySQL database. It appears that the GTDM system can effectively detail an overview process of thesis and dissertation for users.

Following the satisfaction evaluation by 20 users, the result shows that users quite satisfy with the performance of the GTDM system.

Keyword: Enterprise Architecture, Zachman Framework, Graduate Thesis and Dissertation Management System

1. บทนำ

ในปัจจุบันหลักสูตรบัณฑิตศึกษา สาขาฟิสิกส์ ภาควิชาฟิสิกส์ คณะวิทยาศาสตร์ มหาวิทยาลัยมหิดล มีการเก็บรวบรวมข้อมูลต่างๆ มากมาย ทั้งในรูปแบบของกระดาษและไฟล์เอกสารในคอมพิวเตอร์ ทำให้บางครั้งเมื่อต้องการจะค้นหาข้อมูลจึงต้องใช้เวลาในการค้นหาซึ่งค่อนข้างยากในการเข้าถึงข้อมูลที่ต้องการ นอกจากนั้นยังส่งผลกระทบต่อการใช้ข้อมูลมาช่วยในการตัดสินใจ สนับสนุนการบริหาร และการแข่งขันกับคู่แข่งที่ปัจจุบันมีการแข่งขันค่อนข้างสูง ดังนั้นหากสามารถจัดการกับปัญหา อุปสรรคและสามารถสนับสนุนการใช้ข้อมูลได้อย่างเหมาะสมและทันต่อสถานการณ์ก็จะเพิ่มโอกาสการได้เปรียบคู่แข่ง

หน่วยงานอื่นๆ ที่พบปัญหาเกี่ยวกับการจัดการข้อมูล เช่นเดียวกันกับหลักสูตรบัณฑิตศึกษา สาขาฟิสิกส์ และได้นำเทคโนโลยีสารสนเทศ (Information Technology) เข้ามาช่วยจัดการกับข้อมูลเหล่านี้ สามารถจัดการกับข้อมูลต่างๆ ได้อย่างมีประสิทธิภาพ พบว่าไม่ประสบความสำเร็จเสมอไป รวมถึงมีการลงทุนค่อนข้างสูง [4] การนำเทคโนโลยีสารสนเทศเข้ามาใช้จะมีอุปสรรคและความเสี่ยง อย่างไรก็ตาม สถาบันการศึกษา

และหน่วยงานต่างๆ ก็ยังคงมีความพยายามนำเทคโนโลยีสารสนเทศเข้ามาใช้เพราะผลที่ได้ก็นับว่ามีค่าต่ออุปสรรคและความเสี่ยงที่ต้องพบ ขณะเดียวกันก็มีความพยายามในการหาทางแก้ไขปัญหาการนำเทคโนโลยีสารสนเทศเข้ามาใช้วิธีหนึ่ง ที่ประสบความสำเร็จ และได้รับความนิยมในปัจจุบันก็คือการใช้องค์ความรู้ด้านสถาปัตยกรรมการจัดการองค์กร (Enterprise Architecture หรือ EA)

งานวิจัยนี้จะนำโครงร่างแซคแมน (Zachman Framework) ซึ่งเป็น framework ในสถาปัตยกรรมการจัดการองค์กร ร่วมกับภาษายูเอ็มแอล (UML) มาใช้ในการพัฒนาระบบจัดการข้อมูลของหลักสูตรบัณฑิตศึกษาในสาขาฟิสิกส์ มหาวิทยาลัยมหิดล เพื่อลดความเสี่ยง อุปสรรคและความล้มเหลวในการจัดการวิทยานิพนธ์ ของหลักสูตรฯ

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้อง

2.1.1 Zachman framework

Zachman framework [1,2] คือ framework ดั้งเดิมของสถาปัตยกรรมการจัดการองค์กรมีลักษณะรูปแบบเป็นเมทริกซ์ขนาด 6 คูณ 6 โดยในส่วนของคอลัมน์ทั้ง 6 จะประกอบไปด้วยคำถามได้แก่ What, How, Where, Who, When และ Why ในส่วนของแถวทั้ง 6 แถวก็จะเป็นมุมมองจากผู้ที่เกี่ยวข้องกับ framework นั้นๆ อันประกอบไปด้วย Scope, Business model, System model, Technology model, Detailed representation และ Real system

Zachman framework เป็น framework ที่เข้าใจได้ง่าย ไม่มีกระบวนการ รูปแบบ หรือสัญลักษณ์เฉพาะ เพราะเป็นเพียงแต่ตารางธรรมดา ไม่สามารถเพิ่มหรือลดตารางได้ ทำให้จุดแข็งของ framework คือความเป็นตารางธรรมดา ทำให้ครอบคลุมสิ่งสำคัญทั้งหมดที่เราให้ความสนใจและไม่ก่อให้เกิดความซ้ำซ้อน

2.1.2 UML (Unified Modeling Language)

UML [3] คือภาษาที่ได้รับการยอมรับในวงการพัฒนาโปรแกรมและระบบเทคโนโลยีสารสนเทศต่างๆ มีลักษณะเชิงรูปภาพแสดงการทำงานของระบบงาน ในการวิเคราะห์และออกแบบระบบเชิงวัตถุ ทำให้สามารถเข้าใจได้

ง่าย สามารถนำไปใช้ในการพัฒนาโปรแกรมได้อย่างง่ายและรวดเร็ว ปัจจุบันมีถึงเวอร์ชัน UML 2.3

2.2 งานวิจัยที่เกี่ยวข้อง

The Standish Group [4] ทำการสำรวจการนำเทคโนโลยีสารสนเทศมาใช้ภายในประเทศอเมริกา ปรากฏว่าโครงการที่ประสบความสำเร็จ สามารถบรรลุผลได้ภายในเวลาและงบประมาณที่กำหนด รวมถึงสามารถทำงานได้ตามที่ต้องการ ซึ่งพบว่ามีอัตราส่วนเพียงแค่ 16.2% เท่านั้นเมื่อเทียบกับโครงการทั้งหมด ส่วนโครงการที่สำเร็จแต่ใช้เวลาและงบประมาณเกินกว่าที่กำหนดตลอดจนมีการลดความสามารถลงคิดเป็น 52.7% และสุดท้ายมีโครงการที่ต้องยกเลิกไปกลางคันคิดเป็น 31.1%

Zachman [1] ได้ศึกษาและนำเสนอ framework สำหรับช่วยในการนำระบบสารสนเทศเข้ามาใช้กับองค์กรในยุคที่องค์กรมีขนาดใหญ่และความซับซ้อนภายในองค์กรมากขึ้น นับเป็นจุดเริ่มต้นของ Enterprise Architecture ต่อจากนั้น Zachman และ Sowa [2] ได้ช่วยกันรวบรวมข้อมูลที่ผ่านมาจาก framework ที่ได้เสนอไปและเป็นที่ยอมรับของนักวิเคราะห์และออกแบบมาพัฒนาใหม่ในรูปแบบของตาราง Zachman framework ที่ใช้กันอยู่ในปัจจุบัน Roger [5] ได้ทำการเปรียบเทียบ framework ที่ได้รับความนิยม 4 อันดับแรก ซึ่ง Zachman framework ติดอยู่ 4 อันดับแรกและ Zachman framework ยังได้คะแนนในส่วนของการจัดแบ่งประเภท (taxonomy) เป็นอันดับหนึ่งด้วย Dantu และ Smith [6] ยังได้นำ Zachman framework ไปใช้เป็นฉากหลังในการตรวจรักษาทางการแพทย์สำหรับขั้นตอนช่วยในการตัดสินใจ Bahill Botta และ Daniels [7] นำ Zachman framework ไปใช้กับการสร้างโมเดลการเรียนรู้สำหรับกีฬาเบสบอลโดยโมเดลที่ได้ นั้นสามารถทำความเข้าใจได้ง่ายไม่มีความยุ่งยากในการเรียนรู้

Fatolahi และ Shams [3] ได้ทำการศึกษา framework ต่างๆ และพบว่า Zachman framework เป็นหนึ่งในหลาย framework ที่มีอ้างอิงถึงมากที่สุด แต่โมเดลรูปภาพใน Zachman framework ไม่ได้เป็นสิ่งที่ใช้กันอยู่ทั่วไป Fatolahi และ Shams จึงได้นำ UML ซึ่งเป็นโมเดลที่เป็นที่ยอมรับในวงการพัฒนาซอฟต์แวร์ (software) และใช้กันอย่างแพร่หลาย โดยนำมาประยุกต์ใช้กับแถวของ Zachman framework

Hamam และ Loucif [8] ใช้เทคโนโลยี Web-Based เข้ามาช่วยออกแบบและสนับสนุนการตัดสินใจของหลักสูตร ทำให้การเข้าถึงข้อมูลเป็นปัจจุบัน มีความยืดหยุ่นต่อการเข้าถึง และทำให้การทำงานมีประสิทธิภาพมากขึ้น ในงานวิจัยของอัชมา [9] ก็ได้นำเทคโนโลยี Web-Based มาใช้กับการออกแบบและพัฒนาาระบบสารสนเทศเพื่อการจัดการหลักสูตรและทะเบียนรายวิชาในระดับภาควิชาเช่นกัน

จากการศึกษางานวิจัยที่เกี่ยวข้องพบว่ามีการนำ Zachman framework ไปใช้ในองค์กรต่างๆ หลากหลายประเภททั้งบริษัทเอกชน หน่วยงานทางด้านการแพทย์ หรือแม้แต่หน่วยงานด้านกีฬา แต่ยังไม่มีการนำ Zachman framework ไปใช้กับหน่วยงานทางด้านการศึกษา ดังนั้นงานวิจัยนี้จึงเลือก Zachman framework มาใช้กับหน่วยงานทางด้านการศึกษา

3. วิธีการดำเนินการ

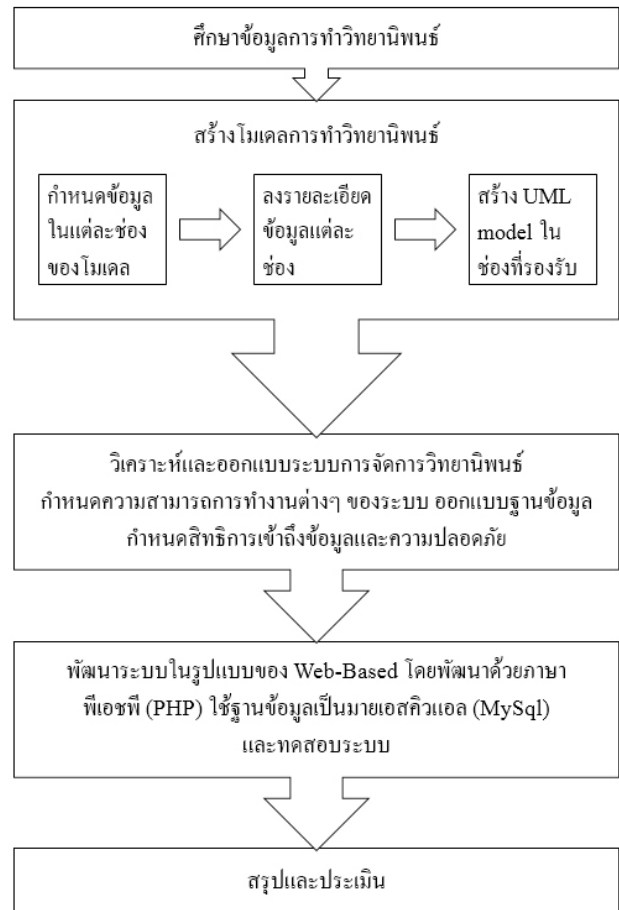
ผู้วิจัยจะทำการศึกษาระบบการทำวิทยานิพนธ์ที่เป็นอยู่ปัจจุบัน เพื่อนำข้อมูลที่ได้มาทำการสร้างโมเดลหลักสูตรบัณฑิตศึกษา จากนั้นจึงนำโมเดลที่ได้ไปวิเคราะห์และออกแบบ เพื่อใช้ในการสร้างระบบการจัดการวิทยานิพนธ์หลักสูตรบัณฑิตศึกษาสำหรับรองรับการจัดการข้อมูลของหลักสูตรฯ โดย Flow chart ในภาพที่ 1 แสดงขั้นตอนการดำเนินงานทั้งหมด

3.1 ศึกษาข้อมูลการทำวิทยานิพนธ์

งานวิจัยนี้จะเริ่มจากศึกษาและรวบรวมข้อมูลการทำวิทยานิพนธ์จากหลักสูตรบัณฑิตศึกษา สาขาฟิสิกส์ มหาวิทยาลัยมหิดล ซึ่งจะประกอบไปด้วย หลักสูตรปรัชญาดุษฎีบัณฑิตสาขาวิชาฟิสิกส์ (หลักสูตรนานาชาติ) และวิทยาศาสตร์มหาบัณฑิตสาขาวิชาฟิสิกส์ (หลักสูตรนานาชาติ) ภาควิชาฟิสิกส์ หลังจากนั้นก็จะนำข้อมูลที่ได้ไปสร้างโมเดล

3.2 สร้างโมเดลการทำวิทยานิพนธ์

โมเดลหลักสูตรบัณฑิตศึกษาจะถูกสร้างตามรูปแบบ Zachman framework ในโมเดลนี้เราจะทำการกำหนดผู้ที่มีส่วนเกี่ยวข้องแต่ละแถวขึ้นมาใหม่ก่อน และในส่วนคอลัมน์จะประกอบไปด้วยคำถามทั้ง 6 ดังแสดงในตารางที่ 1 การกำหนดมุมมองจากผู้ที่เกี่ยวข้องแต่ละแถวเริ่มจากแถวที่ 1 Scope เกี่ยวกับวิสัยทัศน์ พันธกิจ ขอบเขต โครงสร้าง และข้อจำกัด



ภาพที่ 1: Flow chart แสดงขั้นตอนการทำงานทั้งหมด

ตารางที่ 1: ตารางรูปแบบโมเดลการทำวิทยานิพนธ์

	What	How	Where	Who	When	Why
สกอ.						
บัณฑิตศึกษา						
ประธานฯ						
ที่ปรึกษา						
บุคลากร						
นักศึกษา						

ขององค์กรโดยกำหนดให้เป็นสำนักงานคณะกรรมการการอุดมศึกษา (สกอ.) เพราะเป็นผู้กำหนดกฎเกณฑ์ต่างๆ ที่สถาบันการศึกษาระดับอุดมศึกษาต้องปฏิบัติตาม รวมไปถึงเรื่องการทำวิทยานิพนธ์ด้วย แถวที่ 2 Business model เกี่ยวกับเป้าหมาย กลยุทธ์ องค์กรประกอบ และกระบวนการที่ใช้รองรับพันธกิจขององค์กรกำหนดให้เป็นบัณฑิตวิทยาลัย บัณฑิตวิทยาลัยจะเป็นผู้ดูแลหลักสูตรระดับบัณฑิตศึกษาภายในมหาวิทยาลัยทั้งหมดและกำหนดขั้นตอนกฎเกณฑ์ในการทำวิทยานิพนธ์ให้สอดคล้องกับกฎเกณฑ์ของสกอ. แถวที่ 3 System model เกี่ยวกับระบบที่ต้องการ กิจกรรม และขั้นตอน

การทำงานต่างๆ กำหนดให้เป็นประธานหลักสูตรฯ ประธานหลักสูตรฯจะเป็นคนดูแลการทำวิทยานิพนธ์ให้เป็นไปตามขั้นตอนและระเบียบที่บัณฑิตวิทยาลัยกำหนดขึ้น แถวที่ 4 Technology model เกี่ยวกับข้อจำกัดของบุคคล เครื่องมือเทคโนโลยี และวัสดุ กำหนดให้เป็นอาจารย์ที่ปรึกษาและกลุ่มวิจัย แถวที่ 5 Detailed representation เกี่ยวกับบุคคล หรือส่วนประกอบอิสระกำหนดให้เป็นธุรการ ธุรการจะคอยดูแลเอกสารต่างๆ และงานเบื้องหลังอื่นๆ ที่เกี่ยวกับวิทยานิพนธ์ และแถวที่ 6 Real system เกี่ยวกับระบบที่เราต้องการพิจารณา กำหนดให้เป็นนักศึกษา

โดยข้อมูลในแต่ละช่องจะอยู่ในรูปแบบข้อมูลหรือ UML (UML 2.0) ขึ้นอยู่กับความเหมาะสมของข้อมูลนั้นๆ ในงานวิจัยนี้เราจะเริ่มใส่ข้อมูลจากล่างขึ้นบน โดยจะเริ่มจากแถวที่ 6 คอลัมน์ที่ 1 กำหนดให้เป็นวิทยานิพนธ์เพราะเป็นข้อมูลหลักที่สนใจ สัมพันธ์กับคำถาม What และมุมมองนักศึกษาต่อไป แถวที่ 5 คอลัมน์ที่ 1 กำหนดให้เป็นเอกสารและข้อมูลที่เกี่ยวข้องกับวิทยานิพนธ์โดยมีความสัมพันธ์กับวิทยานิพนธ์ที่กำหนดไว้ในแถวที่ 6 คอลัมน์ที่ 1 และสัมพันธ์กับมุมมองของธุรการ คำถาม What การกำหนดข้อมูลลงในช่องจะต้องคำนึงถึงความสัมพันธ์กับแถวและคอลัมน์ และข้อมูลจะต้องมีความสัมพันธ์กับตารางที่อยู่ในแถวเดียวกัน และคอลัมน์เดียวกันทั้งแถว จากนั้นกำหนดแถวถัดไปจนครบทั้งคอลัมน์แล้วจึงไปเริ่มที่คอลัมน์ถัดไปจากแถวล่างสุดคือ แถวที่ 6 คอลัมน์ที่ 2 กำหนดให้เป็นกระบวนการทำวิทยานิพนธ์โดยสัมพันธ์กับคำถาม How มุมมองนักศึกษา และสัมพันธ์กับแถวที่ 6 คอลัมน์ที่ 1 ที่กำหนดให้เป็นวิทยานิพนธ์ด้วย ต่อไปแถวที่ 5 คอลัมน์ที่ 2 ข้อมูลที่จะกำหนดในช่องนี้จะต้องสัมพันธ์กับคำถาม How มุมมองธุรการ และสัมพันธ์กับข้อมูลที่กำหนดไว้แล้วในแถวเดียวกันและคอลัมน์เดียวกัน จากนั้นกำหนดแถวถัดไปจนกว่าจะครบทุกช่อง ดังตารางที่ 2 (ตารางที่ 2 อยู่หน้าสุดท้าย)

สุดท้ายของการสร้างโมเดลคือการนำโมเดลไปให้ผู้ที่มีส่วนเกี่ยวข้องกับระบบตรวจสอบความถูกต้องและความครบถ้วนของข้อมูล

3.3 วิเคราะห์และออกแบบระบบการจัดการวิทยานิพนธ์

จากนั้นเราจะทำการวิเคราะห์ข้อมูลประกอบกับโมเดลที่เราสร้างไว้ มาใช้ในการสร้างระบบการทำวิทยานิพนธ์โดยเริ่ม

จากกำหนดความสามารถการทำงานต่างๆ ของระบบ จากนั้นก็ออกแบบฐานข้อมูล กำหนดสิทธิในการเข้าถึงข้อมูลและความมั่นคงปลอดภัยระบบ

3.4 พัฒนาและทดสอบระบบ

เนื่องจากในปัจจุบันการเข้าถึงอินเทอร์เน็ตเป็นเรื่องที่ง่าย และผู้ใช้งานระบบค่อนข้างหลากหลาย ทำให้เราเลือกที่จะพัฒนาระบบให้อยู่ในรูปแบบของ Web-Based โดยพัฒนาด้วยภาษาพีเอชพี (PHP) และใช้ฐานข้อมูลเป็นมายเอสคิวแอล (MySQL)

โดยหลังจากพัฒนาจะนำระบบไปทดลองใช้งานกับผู้ที่ต้องใช้งานจริงกับระบบ โดยผู้ทดลองมาจากทุกกลุ่มของผู้ที่เกี่ยวข้องกับระบบ หลังจากการทดลองจะทำการปรับปรุงแก้ไขระบบตามข้อมูลที่ได้จากการทดลองใช้งาน

4. ผลการดำเนินการวิจัย

4.1 ผลที่ได้จากการพัฒนาโมเดลการทำวิทยานิพนธ์

โมเดลการทำวิทยานิพนธ์มีรายละเอียด ดังตารางที่ 2 (ตารางที่ 2 อยู่หน้าสุดท้าย) ผลการประเมินจากผู้ที่มีส่วนเกี่ยวข้องกับการทำวิทยานิพนธ์หลังจากได้เห็น โมเดลการทำวิทยานิพนธ์ด้วยวิธีการสัมภาษณ์ โมเดลการทำวิทยานิพนธ์ทำให้เห็นภาพรวมของการทำวิทยานิพนธ์ ตั้งแต่มุมมองของผู้บริหารจนถึงนักศึกษา เห็นข้อมูลที่มีความเกี่ยวข้องกับการทำวิทยานิพนธ์และมีผลต่อการทำวิทยานิพนธ์ ข้อมูลที่แสดงในโมเดลมีการจัดแบ่งประเภท หมวดหมู่ไว้อย่างชัดเจนทำให้โมเดลมีข้อมูลการทำวิทยานิพนธ์ที่ครบถ้วนและเข้าใจได้ง่าย ข้อมูลในบางช่องที่สร้างมาจาก UML ทำให้การวิเคราะห์และออกแบบเป็นไปด้วยความรวดเร็ว ทำให้ได้ Database ที่สมบูรณ์รองรับการใช้งานต่างๆ ที่จะใช้กับระบบดังกล่าวสามารถสร้างระบบที่ใช้งานได้มีประสิทธิภาพ ครบถ้วนตามที่ต้องการ และยังทำให้มองเห็นบางสิ่งบางอย่างที่ยังขาดหรือตกหล่นไป แต่ในทางกลับกัน รายละเอียดแต่ละช่องของโมเดลการทำวิทยานิพนธ์มีค่อนข้างมากทำให้ใช้เวลาในการสร้างที่มากตามไปด้วย

4.2 ผลที่ได้จากการพัฒนาระบบการจัดการวิทยานิพนธ์

ระบบการจัดการวิทยานิพนธ์ทำการแบ่งกลุ่มผู้ใช้งานการเข้าใช้งานระบบ การเข้าถึงข้อมูลและความสามารถในการทำงานของระบบแต่ละผู้ใช้งานจะแตกต่างกันขึ้นอยู่กับสิทธิของผู้ใช้นั้นๆ การดูข้อมูลจะสามารถดูได้เฉพาะข้อมูลของนักศึกษาใน

ความดูแลเท่านั้น และสามารถคัดกรองข้อมูลเบื้องต้นดังภาพที่ 2 หลังจากระบบทำการค้นหาข้อมูลเรียบร้อยแล้ว จะแสดงรายละเอียดข้อมูลดังภาพที่ 3 การได้มาของข้อมูลที่แสดงนี้เป็น การแสดงถึงความสัมพันธ์ข้อมูลจากตารางที่ 2 (ตารางที่ 2 อยู่ หน้าสุดท้าย) ระหว่างนักศึกษากับอาจารย์ในหลักสูตรฯ (ช่องที่ตัดกันระหว่าง Who กับครูกร) นักศึกษาในสังกัด (ช่องที่ตัดกันระหว่าง Who กับที่ปรึกษา, กลุ่มวิจัย) และกระบวนการทำวิทยานิพนธ์ (ช่องที่ตัดกันระหว่าง How กับนักศึกษา)

ภาพที่ 2: ทำการค้นหาข้อมูลที่สนใจ

เรียงลำดับ	ชื่อผลงาน	หลักสูตร	กลุ่มวิจัย	สถานภาพนักศึกษา	สถานะการทำวิทยานิพนธ์	รายละเอียดเพิ่มเติม
1	...	M.Sc.	รายละเอียดเพิ่มเติม
2	...	M.Sc.	รายละเอียดเพิ่มเติม
3	...	M.Sc.	รายละเอียดเพิ่มเติม
4	...	Ph.D.	รายละเอียดเพิ่มเติม
5	...	Ph.D.	รายละเอียดเพิ่มเติม

ภาพที่ 3: ผลการค้นหาข้อมูลเพื่อใช้ในการติดตามดูแลการทำวิทยานิพนธ์

นอกจากนั้นระบบยังสามารถทำการแสดงระยะเวลาเฉลี่ยการทำวิทยานิพนธ์ของนักศึกษาในสังกัด ในแต่ละกลุ่มวิจัย และทั้งหลักสูตรฯ ได้

5. สรุปและข้อเสนอแนะ

งานวิจัยนี้ได้ทำการพัฒนาโมเดลการทำวิทยานิพนธ์และระบบการจัดการวิทยานิพนธ์ หลักสูตรบัณฑิตศึกษาในสาขาฟิสิกส์ มหาวิทยาลัยมหิดล โดยใช้พื้นฐานความรู้ด้าน Enterprise Architecture framework ของ Zachman ร่วมกับภาษายูเอ็มแอล (UML) โมเดลการทำวิทยานิพนธ์ทำให้เห็นภาพรวมของการทำวิทยานิพนธ์และในระหว่างการสร้างโมเดลจะต้องมีการติดต่อสื่อสารกับผู้ที่มีส่วนเกี่ยวข้องกับการทำวิทยานิพนธ์อย่างต่อเนื่องทำให้ได้ข้อมูลที่ครบถ้วน ซึ่งทำให้การพัฒนาการจัดการวิทยานิพนธ์ครอบคลุมและตรงความต้องการของผู้ใช้ โดยการประเมินความพึงพอใจจากกลุ่ม

ตัวอย่างผู้ใช้ระบบการจัดการวิทยานิพนธ์คือ คณาจารย์หลักสูตรบัณฑิตศึกษา สาขาฟิสิกส์ ภาควิชาฟิสิกส์ คณะวิทยาศาสตร์ มหาวิทยาลัยมหิดล จำนวน 19 คน และเจ้าหน้าที่จำนวน 1 คน ได้ผลประเมินค่าความพึงพอใจอยู่ในระดับดี

สำหรับข้อเสนอแนะ ในอนาคตควรมีการพัฒนาให้ระบบการจัดการวิทยานิพนธ์มีความสามารถเพิ่มขึ้น เนื่องจากได้มีการวางรากฐานการพัฒนาไว้อย่างเป็นระบบโดยมีการออกแบบเพื่อรองรับการพัฒนาในอนาคต เช่น ฐานข้อมูลที่มีการออกแบบให้จัดเก็บข้อมูลทั้งระบบไว้แล้ว เป็นต้น และควรนำ Enterprise Architecture มาใช้ร่วมกับการวางแผนการพัฒนาองค์กรด้วยเทคโนโลยีสารสนเทศ ซึ่งปัจจุบันมี Framework ให้เลือกใช้มากมาย

เอกสารอ้างอิง

- [1] John A. Zachman, "A framework for information systems architecture," *IBM Systems Journal* 26, No. 3, 276-292, 1987.
- [2] J. F. Sowa, and J.A. Zachman, "Extending and formalizing the framework for information systems architecture," *IBM Systems Journal* 31, No. 3, 590-616, 1992.
- [3] A. Fatolahi, and F. Shams, "An investigation into applying UML to the Zachman framework," *Inf Syst Front* (2006), 8: 133-143, 2006.
- [4] The Standish Group, "Chaos Report," <http://www.projectsman.com/uk/docs/chaos-report.pdf>, 2007.
- [5] R. Sessions, "A Comparison of the Top Four Enterprise Architecture Methodologies," *MSDN*, 2007.
- [6] B. Dantu, and E. Smith, "Medical Process Modeling with a Hybrid System Dynamics Zachman Framework," *Procedia Computer Science* 6 (2011), 76-81, 2011.
- [7] T. Bahill, R. Botta, and J. Daniels, "The Zachman Framework Populated with Baseball Models," *Journal of Enterprise Architecture*, November 2006, 2006.
- [8] H. Hamam, and S. Loucif, "Web-Based Engine for Program Curriculum Designers," *IEEE TRANSACTIONS ON EDUCATION*, VOL. 52, NO. 4, NOVEMBER 2009, 2009.
- [9] อัจฉา มุจลินท์พิงกูร, "การออกแบบและพัฒนาระบบสารสนเทศเพื่อการจัดการหลักสูตรและทะเบียนรายวิชาในระดับภาควิชา".

วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2006.

ตารางที่ 2: โมเดลการทำวิทยานิพนธ์

	What	How	Where	Who	When	Why
สกอ.	กฎที่เกี่ยวกับวิทยานิพนธ์	กระบวนการต่างๆ ที่เกี่ยวข้อง	ประเทศไทย	สถาบันระดับอุดมศึกษาในไทย	เปิดหลักสูตรระดับบัณฑิตศึกษา	วิสัยทัศน์ พันธกิจของสกอ.
บัณฑิตศึกษา	ข้อกำหนดวิทยานิพนธ์	กระบวนการทำวิทยานิพนธ์	มหาวิทยาลัยมหิดล	สถาบันภาควิชา หน่วยในมหิดล	เปิดหลักสูตรระดับบัณฑิตศึกษา	วิทยานิพนธ์ตรงตามกฎของสกอ.
ประธานฯ	ข้อมูลการทำวิทยานิพนธ์	ดูแลติดตามการทำวิทยานิพนธ์	ภาควิชาฟิสิกส์	นักศึกษาบุคลากรในหลักสูตรฯ	มีนักศึกษาทำวิทยานิพนธ์ในหลักสูตรฯ	หลักสูตรฯ ประสบความสำเร็จ
ที่ปรึกษา, กลุ่มวิจัย	ข้อมูลต่างๆ ในการทำ	ให้คำปรึกษาฝึกการทำ	สถานที่ทำวิจัยในสังกัด	นักศึกษา และเครื่องมือในสังกัด	นักศึกษาขอคำปรึกษา รายงานผล ฯ	ได้วิทยานิพนธ์ที่ดี นักศึกษาใช้เวลาในกำหนด
ธุรการ	เอกสารและข้อมูลที่เกี่ยวข้องวิทยานิพนธ์	การทำหนังสืองานธุรการ	ธุรการฟิสิกส์บัณฑิตพัฒนาไต่บัณฑิตศาลายา	นักศึกษ้อาจารย์ในหลักสูตรฯ	นักศึกษายื่นคำร้อง ขอความช่วยเหลือ	การทำวิทยานิพนธ์ราบรื่น
นักศึกษา	วิทยานิพนธ์	กระบวนการทำวิทยานิพนธ์	ห้องทำวิจัย	นักศึกษาลงทะเบียนวิทยานิพนธ์	ลงทะเบียนวิทยานิพนธ์	สำเร็จการศึกษา

ระบบติดตามและตรวจสอบย้อนกลับก้อนเชื้อเห็ด Mushrooms Cube Tracking and Traceability System

อนุวัฒน์ ใจดี (Anuwat Jaidee)¹ และบุษยดี ศิริแสงตระกูล (Pusadee Seresangtakul)²

^{1,2}ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น
anuwat_jaidee@glive.kku.ac.th¹, pusadee@kku.ac.th²

บทคัดย่อ

ปัจจุบันผู้บริโภคต่างตระหนักถึงความปลอดภัยจากการบริโภคผลิตภัณฑ์เกษตร เห็ดเป็นหนึ่งในอาหารที่คนไทยนิยมบริโภค จากสภาพแวดล้อมและกระบวนการผลิตที่เปลี่ยนแปลงทำให้ผู้บริโภคสนใจข้อมูลเกี่ยวกับผลิตภัณฑ์และแหล่งที่มาของสินค้ามากขึ้น งานวิจัยนี้จึงนำเสนอระบบตรวจสอบย้อนกลับในกระบวนการผลิตก้อนเชื้อเห็ดตั้งแต่แหล่งวัตถุดิบ แหล่งผลิตจนถึงผู้บริโภค ซึ่งประกอบด้วย ระบบจัดการข้อมูลฟาร์มเห็ด และ ระบบตรวจสอบย้อนกลับก้อนเชื้อเห็ดด้วยการบูรณาการเทคโนโลยีอาร์เอฟไอดีและบาร์โค้ด ซึ่งระบบที่นำเสนอจะเป็นการช่วยส่งเสริมธุรกิจด้านการเกษตรของไทยให้เข้าสู่มาตรฐานสากล

คำสำคัญ: อาร์เอฟไอดี ก้อนเชื้อเห็ด บาร์โค้ด การตรวจสอบย้อนกลับ

Abstract

Nowadays, people have focused on the quality and safety of food or products from agricultural industries. Mushroom is one of popular food in Thailand. Due to environment and production process changed, consumers are aware of agricultural food safety and concern about food information. This work proposed a traceability system for mushroom cubes from supplier, producer to consumer by using RFID and barcode technologies. In order to track and trace the information, the mushroom farm management and traceability sub-systems were developed. The proposed system will lead the agricultural product to international standard.

Keyword: RFID, Mushrooms-cube, Barcode, Traceability

1. บทนำ

ปัจจุบันหลายประเทศให้ความสำคัญกับคุณภาพและความปลอดภัยในสินค้าและอาหารจากอุตสาหกรรมเกษตร เพราะผู้บริโภคได้ตระหนักถึงผลกระทบที่เกิดจากการบริโภคสินค้ามากขึ้น อีกทั้งในสหภาพยุโรป สหรัฐอเมริกาและญี่ปุ่น ได้มีกฎหมายเพื่อกำหนดมาตรฐานสินค้าทางการเกษตรและประกาศและบังคับใช้อย่างเป็นทางการ เพื่อให้สามารถสอบย้อนกลับสินค้าเกษตรถึงแหล่งที่มา [5] สำหรับประเทศไทย สำนักงานมาตรฐานสินค้าเกษตรและอาหารแห่งชาติ (มกอช.) กระทรวงเกษตรและสหกรณ์ได้ผลักดันให้สินค้าเกษตรของไทยมีระบบการสอบย้อนกลับ เพื่อความสามารถในการส่งออกระหว่างประเทศ ซึ่งสินค้าที่ได้รับรองมาตรฐานต้องมีข้อมูลผู้ผลิตและผู้จัดจำหน่าย ถ้าเป็นสินค้านำเข้าต้องระบุที่อยู่และที่ตั้งของผู้นำเข้า รวมถึงข้อมูลแหล่งผลิตโดยให้ระบุประเทศผู้ผลิตและพื้นที่ผลิต [1]

เห็ดเป็นหนึ่งในสินค้าเกษตรของไทยที่เป็นที่ต้องการของตลาด [9] เนื่องจากเป็นแหล่งโปรตีนที่มีระยะเวลาการผลิตค่อนข้างสั้น ปัจจุบันเกษตรกรผู้ผลิตก้อนเชื้อเห็ดยังขาดการรับรองมาตรฐานคุณภาพสินค้าและมาตรฐานระบบ ทำให้ประสบกับปัญหาในการตรวจสอบย้อนกลับคุณภาพก้อนเชื้อเห็ดในส่วนผู้ผลิตก้อนเชื้อเห็ดไปยังลูกค้าหรือฟาร์มเปิดดอกเห็ด รวมถึงผู้บริโภค ดังนั้นงานวิจัยนี้ จึงนำเสนอระบบสอบย้อนกลับข้อมูลในกระบวนการผลิตก้อนเชื้อเห็ด ด้วยการประยุกต์ใช้สถาปัตยกรรมอาร์เอฟไอดี และบาร์โค้ด เพื่อเป็นประโยชน์ในการตรวจสอบย้อนกลับถึงแหล่งที่มาของก้อนเชื้อเห็ด และวัตถุดิบที่ใช้ในการผลิต

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ระบบการสอบย้อนกลับ

องค์กร GS1 (The Global GS1 Traceability Standard) เดิมคือ EAN International ซึ่งเป็นองค์กรระดับชาติ ร่วมกับ Uniform Code Council (UCC) USA กำหนดและผลักดันให้เกิดมาตรฐานการสอบย้อนกลับสากลรวมถึงวางมาตรฐานของการจัดเก็บข้อมูลอัตโนมัติ และการสื่อสารอิเล็กทรอนิกส์ในอุตสาหกรรม ปัจจุบันมีผู้ใช้ทั่วโลก รวมถึงประเทศไทย กระบวนการสอบย้อนกลับประกอบด้วย 2 ขั้นตอน [4] คือ

1) ค้นหาแหล่งที่มาของสินค้า ถ้าพบสินค้ามีปัญหาที่ปลายทาง (ผู้บริโภครหรือลูกค้า) จำเป็นต้องตรวจสอบย้อนกลับถึงต้นน้ำ (แหล่งผลิตหรือที่มาของวัตถุดิบ) เพื่อหาสาเหตุของปัญหาว่าเกิดขึ้นที่ขั้นตอนใด เช่น วัตถุดิบ การผลิต การจัดเก็บ การขนส่ง เป็นต้น

2) การติดตามที่จัดส่งสินค้า เพื่อเรียกคืนสินค้าในกรณีตรวจพบว่ากระบวนการผลิตบางล๊อตมีปัญหา ซึ่งอาจเป็นอันตรายต่อลูกค้าหรือผู้บริโภค

2.2 เทคโนโลยีอาร์เอฟไอดี RFID และบาร์โค้ด Barcode

อาร์เอฟไอดี (Radio Frequency Identification: RFID) [3] เป็นระบบระบุลักษณะวัตถุด้วยคลื่นความถี่วิทยุมีจุดเด่นคืออ่านข้อมูลจากแท็ก (Tag) ได้หลายแท็กแบบไร้สัมผัสด้วยความเร็วสูง โดยข้อมูลจะถูกเก็บไว้ในไมโครชิปที่อยู่ในแท็ก

บาร์โค้ด [3] คือสัญลักษณ์รหัสแท่งที่ใช้แทนข้อมูลตัวเลขหรือตัวอักษร โดยประกอบด้วยแท่งบาร์สีเข้มและช่องว่างสีอ่อนเรียงสลับกัน มาตรฐานบาร์โค้ดที่ใช้กันมากคือ EAN (European Article Number) ซึ่งกำหนดมาตรฐานโดย EAN International และ UPC (Universals Product Code) สามารถแบ่งชนิดบาร์โค้ด ดังนี้

1) GTIN-13 เป็นบาร์โค้ดที่ใช้ติดบนสินค้าสำหรับขายปลีกทั่วไป

2) GTIN-14 เป็นบาร์โค้ดใช้สำหรับสินค้าค้าส่ง โดยสามารถบอกถึงความแตกต่างของหน่วยบรรจุภัณฑ์

3) GS1-128 เป็นบาร์โค้ดที่สามารถเก็บข้อมูลได้ทั้งตัวเลขและตัวอักษร จึงสามารถบันทึกข้อมูลของสินค้าได้มากสามารถนำไปใช้ประโยชน์ในการบริหารจัดการด้านโลจิสติกส์และซัพพลายเชนและระบบการสืบค้นย้อนกลับ (Traceability)

ได้อย่างมีประสิทธิภาพ ดังนั้นงานวิจัยนี้จึงเลือกใช้บาร์โค้ดแบบ GS1-128 จึงสามารถเก็บข้อมูลได้มากกว่าแบบอื่น

2.3 งานวิจัยที่เกี่ยวข้อง

จากการศึกษางานวิจัยที่เกี่ยวข้องพบว่า มีการนำเทคโนโลยีอาร์เอฟไอดี มาประยุกต์ใช้ในระบบการสอบย้อนกลับข้อมูลในสินค้าด้านการเกษตร เช่น วรพจน์ จักขุพันธ์ และคณะ[5] ได้วิเคราะห์การไหลของข้อมูลในห่วงโซ่อุปทานการผลิตและตรวจสอบย้อนกลับการผลิตเนื้อหมู เลือกจุดควบคุม การเก็บรวบรวม การส่งผ่านข้อมูล รวมถึงออกแบบซอฟต์แวร์ให้มีการเก็บข้อมูลและส่งข้อมูลอัตโนมัติบนพื้นฐานของเทคโนโลยีอาร์เอฟไอดี

อภิวิชญ์ และคณะ[2] ได้เสนอการบริหารจัดการโลจิสติกส์ภายในโรงงานแป้งมันสำปะหลังด้วย RFID และการเชื่อมโยงข้อมูลเข้ากับระบบสารสนเทศภายในโรงงานด้วยเว็บเซอร์วิสเพื่อลดต้นทุนทางโลจิสติกส์และสร้างความได้เปรียบกับคู่แข่ง

สมจิตร และคณะ[8] ได้ออกแบบซอฟต์แวร์เพื่อการสอบย้อนกลับข้าวหอมมะลิในห่วงโซ่อุปทานข้าวหอมมะลิระบบประกอบด้วย 4 ส่วน คือ การแลกเปลี่ยนเอกสารอิเล็กทรอนิกส์ การออกแบบรหัสมาตรฐาน ออกแบบการสอบย้อนกลับ การออกแบบซอฟต์แวร์ตามมาตรฐาน ทำให้ได้ซอฟต์แวร์ตามมาตรฐานสามารถบูรณาการเข้ากับระบบอื่นได้อย่างมีประสิทธิภาพ

Yu-Chia Hsu, An-Pin Chen and Chun-Hung Wang [10] ได้นำเสนอระบบตรวจสอบย้อนกลับห่วงโซ่อุปทานของปลา โดยการประยุกต์ใช้อาร์เอฟไอดี ห่วงโซ่อุปทานของปลาสดประกอบไปด้วยฟาร์มเลี้ยงปลา ผู้ตรวจสอบ ศูนย์โลจิสติกส์ และร้านอาหาร

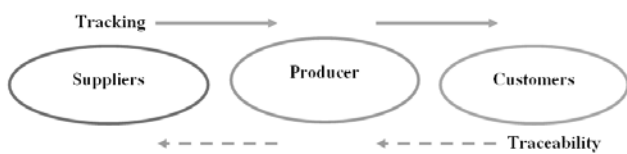
Ren Shougang, Xu Huanliang, Lian and Zhou Guanghong [7] นำเสนอระบบติดตามและสอบย้อนกลับผลิตภัณฑ์เนื้อสัตว์ด้วยเทคโนโลยี RFID จะเก็บข้อมูลในวงจรการผลิตทั้งที่เกี่ยวข้องกับจุดจำหน่ายและตลอดจนตรวจสอบย้อนกลับจากลูกค้าไปยังจุดจำหน่ายหรือค้าปลีก และสุดท้ายที่ผู้ผลิตหรือโรงงาน

งานวิจัยนี้จึงได้ประยุกต์ใช้อาร์เอฟไอดีและบาร์โค้ดเพื่อการติดตามและตรวจสอบย้อนกลับข้อมูลในธุรกิจเนื้อ

3. การวิเคราะห์และออกแบบระบบ

3.1 ห่วงโซ่อุปทานและกระบวนการผลิตก่อนเชื้อเห็ด

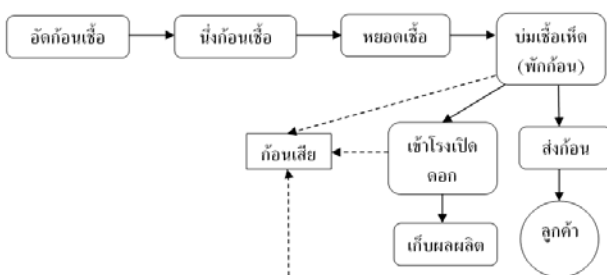
ในงานวิจัยนี้ได้เน้นทำการศึกษากระบวนการผลิตก่อนเชื้อเห็ดของฟาร์มเห็ดใหม่เจริญทรัพย์ อำเภอน้ำพอง จังหวัดขอนแก่นซึ่งเป็นฟาร์มเห็ดขนาดกลางที่ผลิตผลิตภัณฑ์เกี่ยวกับเห็ดครบวงจร สามารถแสดงห่วงโซ่อุปทานของระบบได้ดังภาพที่ 1 ซึ่งประกอบด้วย ผู้จัดหาวัตถุดิบ (Suppliers) ฟาร์มผู้ผลิตก่อนเชื้อเห็ด (Producer) และฟาร์มลูกค้า (Customers) ตั้งแต่ต้นน้ำจนถึงปลายทางตามลำดับ



ภาพที่ 1: ห่วงโซ่อุปทานการผลิตก่อนเชื้อเห็ด

จากการศึกษาระบบงานของฟาร์มใหม่เจริญทรัพย์ ประกอบด้วยระบบงานหลัก 3 ส่วนดังนี้ 1) ระบบจัดซื้อวัตถุดิบในการผลิต 2) ระบบการผลิตก่อนเชื้อเห็ด การผลิตดอกเห็ด และผลิตอุปกรณ์เพื่อใช้ในการเห็ด 3) ระบบจัดจำหน่ายผลิตภัณฑ์

ในส่วนของกระบวนการผลิต วัตถุดิบที่นำเข้าจะถูกนำเข้าสู่กระบวนการผลิตก่อนเชื้อเห็ดเพื่อจัดส่งให้กับลูกค้า ดังแสดงกระบวนการทำงานในภาพที่ 2

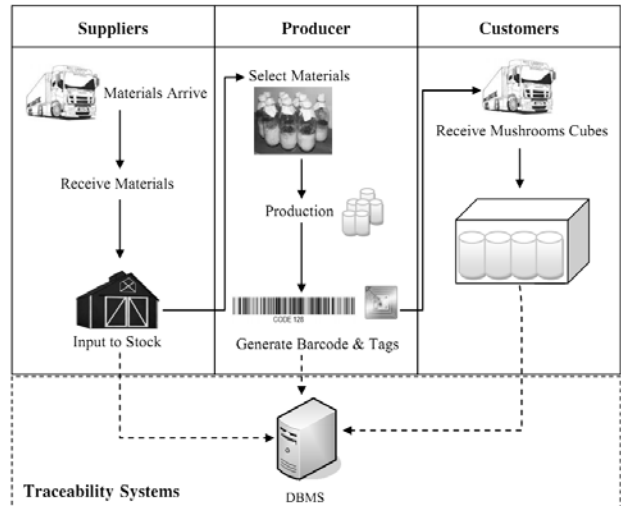


ภาพที่ 2: กระบวนการผลิตก่อนเชื้อเห็ด

เพื่อให้ผู้บริโภคสามารถตรวจสอบย้อนกลับข้อมูลถึงแหล่งที่มาของวัตถุดิบและข้อมูลการผลิตได้ ในการออกแบบจึงต้องมีการบันทึกข้อมูลทุกขั้นตอนในห่วงโซ่อุปทาน ซึ่งจาก

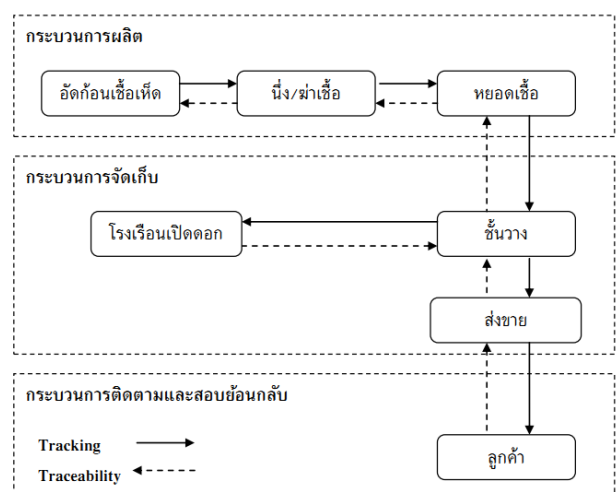
การวิเคราะห์ สามารถออกแบบระบบและแสดงการไหลของข้อมูลได้ดังภาพที่ 3 ซึ่งจะประกอบด้วย 3 ส่วนคือ

1) ผู้จัดหาวัตถุดิบ (Suppliers) จะทำการจัดหาวัตถุดิบ เช่น ขี้เลื่อย เชื้อเห็ด ปูนขาว ภูไมท์ กากน้ำตาล ยิปซัม ฯลฯ และอุปกรณ์การผลิต เพื่อส่งมายังผู้ผลิตและจัดเก็บลงสต็อก



ภาพที่ 3: การไหลข้อมูลของระบบ

2) ผู้ผลิต (Producer) เลือกว่าวัตถุดิบและอุปกรณ์การผลิตจำนวนที่ผลิต เมื่อเข้าสู่กระบวนการผลิต ซึ่งในกระบวนการจะมีการบันทึกข้อมูลแต่ละขั้นตอนเพื่อให้สามารถสอบย้อนกลับได้ (ดังแสดงในภาพที่ 4)



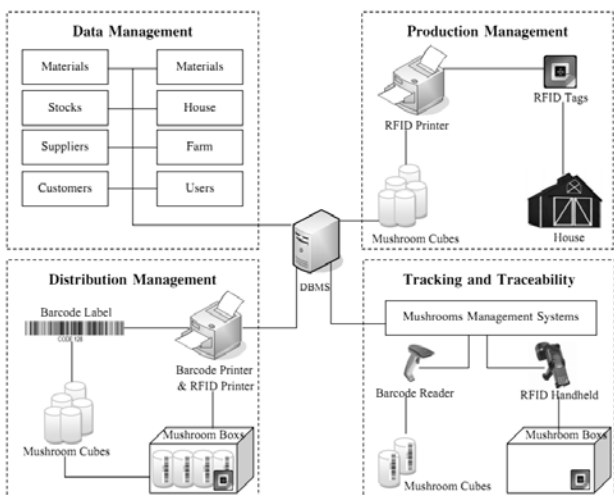
ภาพที่ 4: ออกแบบระบบการผลิตก่อนเชื้อเห็ด

3) ลูกค้า (Customers) เมื่อก่อนซื้อเห็ดถึงฟาร์มลูกค้า จะมีกระบวนการตรวจรับก้อนเชื้อเห็ดโดยอ่านแท็กจากถังบรรจุก้อนเชื้อเห็ด เพื่อตรวจสอบความถูกต้อง และครบถ้วนของก้อนเชื้อเห็ด

แต่ละส่วนที่เกี่ยวข้อง อาทิเช่น ผู้จัดการฟาร์มผู้ผลิต และลูกค้า จะจัดเก็บข้อมูลทุกกระบวนการเข้าฐานข้อมูลส่วนกลางของระบบฯ เพื่อง่ายต่อการตรวจสอบย้อนกลับในกระบวนการผลิตได้

3.2 การออกแบบสถาปัตยกรรมของระบบ

การออกแบบระบบสอบย้อนกลับของก้อนเชื้อเห็ดเพื่อให้สามารถสอบย้อนกลับข้อมูลได้ ระบบที่นำเสนอประกอบด้วย 2 ส่วน ได้แก่ ระบบบริหารจัดการฟาร์มเห็ด และ ระบบสอบย้อนกลับข้อมูล ภาพที่ 5 แสดงสถาปัตยกรรมของระบบ โดยสามารถประยุกต์ใช้ในแต่ละขั้นตอนของกระบวนการผลิตดังแสดงในภาพที่ 6



ภาพที่ 5: สถาปัตยกรรมระบบสอบย้อนกลับก้อนเชื้อเห็ด

จากการศึกษาวิเคราะห์ระบบงานเดิมสามารถออกแบบระบบบริหารจัดการฟาร์มเห็ด ประกอบด้วย 3 โมดูลคือ โมดูลการจัดการข้อมูล (Data Management) โมดูลจัดการถือการผลิต (Product management) และโมดูลกระจายสินค้า

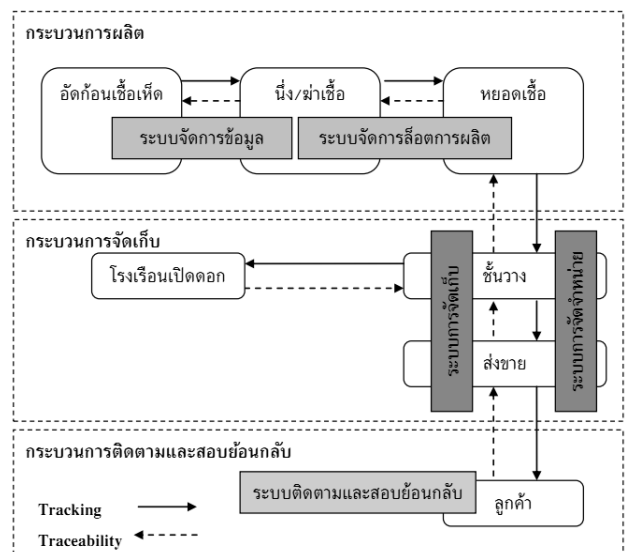
1) โมดูลการจัดการข้อมูล ประกอบด้วย 4 โมดูลย่อย

1.1) ระบบจัดการข้อมูล จะจัดการข้อมูลพื้นฐาน เช่น ข้อมูล วัตถุดิบ สต็อกวัตถุดิบ ข้อมูลลูกค้า ข้อมูลตัวแทนจำหน่าย ข้อมูลซัพพลายเออร์

1.2) ระบบจัดการถือการผลิต ประกอบด้วย การสร้างหมายเลขบาร์โค้ดประจำตัวก้อนเชื้อเห็ด หมายเลขถือการผลิต จำนวนการผลิต วันที่ผลิต แหล่งที่มาของวัตถุดิบ

1.3) ระบบการจัดเก็บ คือบันทึกข้อมูลการจัดเก็บก้อนเชื้อเห็ด เช่น หมายเลขแท็กโรงเรือน หมายเลขแท็กชั้นวาง หมายเลขแท็กถังบรรจุ

1.4) ระบบการจัดจำหน่าย บันทึกการขายขายจนกระทั่งถึงกระบวนการจัดส่งถึงลูกค้าอย่างครบถ้วนสมบูรณ์



ภาพที่ 6: ประยุกต์โมดูลเข้ากับกระบวนการผลิต

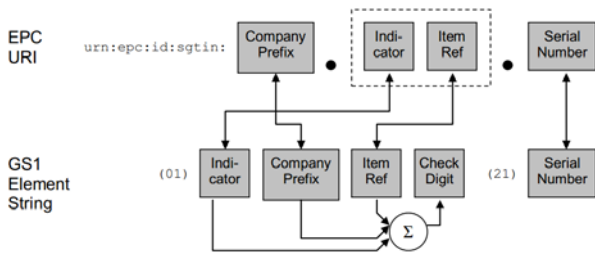
2) จัดการถือการผลิต จะทำการจัดการข้อมูลแท็กอาร์เอฟไอดี และบาร์โค้ด

2.1) การออกแบบรหัสสินค้า

เพื่อให้ระบบเป็นไปตามมาตรฐานสากล ในงานวิจัยนี้ได้ทำการ ออกแบบหมายเลขแหล่งที่ตั้งของฟาร์ม ตามที่สำนักมาตรฐานสินค้าเกษตรและอาหารแห่งชาติ(มกอช.) ได้นำร่องการใช้เทคโนโลยีเพื่อการตรวจสอบแหล่งผลิตสินค้าเกษตรและอาหารตามมาตรฐานสากล โดยใช้รหัส SGLN (Serial Global Location Number) [6] เพื่อให้เป็นมาตรฐานเดียวกัน

2.2) ออกแบบหมายเลขประจำตัวก้อนเชื้อเห็ดและถังเห็ดจะออกแบบตามมาตรฐานของอีพีซี SGTIN-96 [6] (โครงสร้างดังปรากฏในภาพที่ 7) เพื่อรองรับการใช้งานร่วมกันทั้งบาร์โค้ดและอาร์เอฟไอดีตามมาตรฐาน GS1 และสอดคล้องกับมาตรฐานของ มกอช. ซึ่งมีรายละเอียดดังนี้

- Company Prefix แสตรหัสของบริษัทหรือองค์กร โดยค่าจะไม่ซ้ำกันเพื่อให้สามารถระบุได้ว่าสินค้าที่ได้รับการติดแท็กมาจากบริษัทไหน รหัสจะถูกจัดสรรโดยองค์กรไอเอเอ็น (EAN) หรือยูซีซี (UCC) ซึ่งต่อมารวมกันเป็นจีเอสวัน (GS1) ในประเทศนั้น ๆ สำหรับประเทศไทยสามารถขอรหัสนี้ได้จาก สถาบันรหัสสากล (GS1 Thailand) สภาอุตสาหกรรมแห่งประเทศไทย



ภาพที่ 7: ส่วนประกอบของ SGTIN-96 [6]

- Item Reference และ Indicator Digit เป็นส่วนที่ระบุกรณีที่มีศูนย์นำ เช่น “00123” จะต่างจาก “123” โดยจะใช้ส่วนบ่งชี้ที่มีค่าเท่ากับ 1 ใส่เพิ่มเข้าไปเป็นตัวหน้าสุด ซึ่งจะได้ผลลัพธ์ดังนี้ “100123” เพื่อใช้เป็นตัวระบุความแตกต่างจาก “123” สำหรับการใส่ส่วนบ่งชี้ (Indicator Digit) ที่เท่ากับ 1 นั้น จะทำในเลขฐานสิบแล้วจึงแปลงเป็นเลขฐานสองก่อนเข้าไปเก็บในส่วนอ้างอิง (Item Reference)

- Serial Number ส่วนของการจัดเก็บข้อมูลเกี่ยวกับผลิตภัณฑ์แบ่งบรรจุ หรือผลิตภัณฑ์เฉพาะภาชนะบรรจุภัณฑ์อย่างเดียวโดยอีพีซีโกบอลจะจัดการให้ค่าต่างของตัวเลขแต่ละคลาสวัตถุ (Object Class) มีค่าไม่ซ้ำกัน [7]

ถ้าให้หมายเลขรหัสของบริษัท (Company Prefix) เป็น 0614141 และหมายเลขรหัสส่วนอ้างอิง (Item Reference) เป็น 056478 และ Serial Number เป็น 1 จะได้ดังภาพที่ 8

ระบบสอบย้อนกลับข้อมูลก่อนซื้อเห็ด แบ่งออกเป็น 4 ส่วน ดังนี้

- 1) การจัดการข้อมูล ประกอบด้วย ข้อมูลวัตถุดิบ สัตว์กผู้จัดหาวัตถุดิบ ลูกค้า โรงเรือน ฟาร์ม และจัดการผู้ใช้งานระบบ
- 2) การจัดการข้อมูลการผลิต เมื่อการผลิตแต่ละล็อตเสร็จสมบูรณ์ ระบบจะสร้างหมายเลขบาร์โค้ดประจำก้อนเชื้อเห็ด



ภาพที่ 8: การออกแบบหมายเลขประจำตัวก้อนเชื้อเห็ด

3) การจัดและสร้างหมายเลขแท็กเพื่อติดที่โรงเรือนหรือชั้นวางที่จะนำก้อนเชื้อเห็ดไปเก็บหรือบ่มเพาะเชื้อ เมื่อมีลูกค้าสั่งก้อนเชื้อเห็ดเข้ามา ผู้ผลิตจะจัดส่งก้อนเชื้อเห็ดให้ลูกค้า โดยติดสติ๊กเกอร์บาร์โค้ดที่ก้อนเชื้อเห็ดแล้วบรรจุในลังบรรจุเห็ดที่มีแท็กอาร์เอฟไอดีอยู่ 1 ลังบรรจุก้อน เชื้อเห็ด 12 ก้อน เมื่อขนส่งก้อนเชื้อเห็ดถึงฟาร์มลูกค้าแล้ว จะใช้อาร์เอฟไอดีมือถืออ่านข้อมูลแท็กที่ลังบรรจุเห็ดเพื่อเก็บข้อมูลการจัดส่งเข้าสู่ระบบ

4) การติดตามและสอบย้อนกลับก้อนเชื้อเห็ดสามารถสอบย้อนกลับได้ 2 รูปแบบ คือ สอบย้อนกลับออนไลน์ผ่านหน้าเว็บไซต์ของผู้ผลิต โดยอ่านบาร์โค้ดจากก้อนเชื้อเห็ด และสอบย้อนกลับแบบออฟไลน์ ซึ่งจะใช้อาร์เอฟไอดีมือถืออ่านข้อมูลแท็กที่ลังบรรจุเห็ดหรืออ่านบาร์โค้ดที่ก้อนเชื้อเห็ด แล้วจัดเก็บเป็นเท็กซ์ไฟล์ *.txt เพื่ออัปโหลดเข้าสู่ระบบในภายหลัง

4. การพัฒนาระบบ

ส่วนติดต่อผู้ใช้งานเป็นเว็บแอปพลิเคชัน พัฒนาด้วยภาษา PHP จัดเก็บข้อมูลด้วยระบบจัดการฐานข้อมูลเชิงสัมพันธ์ MySQL มีกระบวนการทำงาน 5 ส่วน คือ

- 1) ส่วนจัดการข้อมูลพื้นฐาน เช่น ข้อมูลฟาร์ม ข้อมูลโรงเรือนและชั้นวาง ข้อมูลวัตถุดิบ ข้อมูลเห็ด ข้อมูลแหล่งที่มาของวัตถุดิบ ข้อมูลลูกค้า
- 2) ส่วนรับเข้าข้อมูลวัตถุดิบในการผลิต
- 3) ส่วนจัดการถือการผลิต ซึ่งจะใช้ข้อมูลพื้นฐาน ข้อมูลสัตว์กวัตถุดิบ และข้อมูลโรงเรือนและชั้นวาง มาสร้างเป็นข้อมูลถือการผลิต

- 4) ส่วนการจัดเก็บก้อนเชื้อเห็ด
- 5) ส่วนการจัดจำหน่ายก้อนเชื้อเห็ด

จากการพัฒนาระบบจัดการข้อมูลฟาร์มเห็ด ผู้ใช้สามารถจัดการข้อมูลพื้นฐาน ข้อมูลแหล่งที่มาวัตถุดิบ ข้อมูลลูกค้า ลีตการผลิต และข้อมูลการจัดจำหน่ายได้ ดังภาพที่ 9



ภาพที่ 9: ระบบจัดการข้อมูลฟาร์มเห็ด

5. บทสรุป

งานวิจัยนี้ได้นำเสนอกระบวนการออกแบบระบบติดตามและตรวจสอบย้อนกลับก้อนเชื้อเห็ด เพื่อยกระดับมาตรฐานสินค้าเกษตรของไทยตามข้อบังคับของสำนักงานมาตรฐานสินค้าเกษตรและอาหารแห่งชาติ (มกอช.) กระทรวงเกษตรและสหกรณ์ โดยผู้วิจัยได้ออกแบบกระบวนการผลิตก้อนเชื้อเห็ดและประยุกต์ใช้เทคโนโลยีบาร์โค้ดและเทคโนโลยีอาร์เอฟไอดี ออกแบบการรับส่งข้อมูลระหว่างองค์กร และการออกแบบรหัสมาตรฐานประจำตัวก้อนเชื้อเห็ดเพื่อรองรับการสืบย้อนกลับ โดยอ้างอิงตามมาตรฐานทำให้ระบบสามารถบูรณาการเข้ากับระบบหรือมาตรฐานอื่นได้ง่ายขึ้น ไม่ว่าจะอยู่ต่างแพลตฟอร์มก็จะสามารถเชื่อมโยงข้อมูลได้อย่างมีประสิทธิภาพ

แนวทางในการดำเนินงานต่อไป นำต้นแบบไปติดตั้งและทดสอบกับระบบงานจริง เพื่อประเมินประสิทธิภาพต่อไป นอกจากนี้ ข้อมูลที่ได้จากการจัดเก็บสามารถนำไปใช้ในกระบวนการตัดสินใจได้

เอกสารอ้างอิง

- [1] คณะกรรมการมาตรฐานสินค้าเกษตรและอาหารแห่งชาติ, “ประกาศคณะกรรมการมาตรฐานสินค้าเกษตรและอาหารแห่งชาติ เรื่อง กำหนดมาตรฐานสินค้าเกษตรและอาหารแห่งชาติ: เห็ดหอมสด พ.ศ. 2541,” *ราชกิจจานุเบกษา*, เล่ม 125 ตอนพิเศษ 139 ง, ลงวันที่ 18 สิงหาคม 2551.
- [2] อภิวิชญ์ เจษฎาพันธุ์ และ สมจิตร อาจอินทร์, “ระบบจัดการโลจิสติกส์ ภายในโรงงานแปงมันสำปะหลังด้วย RFID กรณีศึกษา บริษัท สงวนวงษ์อุตสาหกรรม จำกัด,” *การประชุมสัมมนาวิชาการด้านการจัดการโลจิสติกส์และโซ่อุปทาน ครั้งที่ 9*, 2552, หน้า 13-23.
- [3] สถาบันส่งเสริมความเป็นเลิศทางเทคโนโลยีอาร์เอฟไอดีแห่งประเทศไทย, “แนะนำเทคโนโลยีอาร์เอฟไอดี,” [ออนไลน์] 2555. จาก: <http://www.rfid.or.th/2012/th/technology/know.asp> [1 มีนาคม 2555].
- [4] เฉลิมชนม์ ไชยคำตรง, “ระบบมาตรฐานการสอบย้อนกลับสากล,” *สำนักโลจิสติกส์อุตสาหกรรม กรมอุตสาหกรรมพื้นฐานและการเหมืองแร่*, 2554, หน้า 1-6.
- [5] วรพจน์ จักขุพันธ์, “ระบบสืบย้อนกลับโดยอาศัย RFID และมาตรฐานรหัสสากล กรณีศึกษา: การสืบย้อนกลับเนื้อสุกรจากร้านค้าปลีกสุฟาร์ม,” [วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ]. ขอนแก่น: บัณฑิตวิทยาลัยมหาวิทยาลัยขอนแก่น; 2550.
- [6] EPCglobal, “EPCTM Generation 1 Tag Data Standards Version 1.1 Rev.1.27,” *Standard Specification*, 10 May 2005, pp.1-87.
- [7] Ren Shougang, Xu Huanliang, Lian and Zhou Guanghong, “Research on RFID-Based Meat Product Track and Traceability System,” *International Conference on Computer Application and System Modeling (ICCSM)*, Taiyuan, 22-24 Oct. 2010, Vol.8, pp.458-462.
- [8] สมจิตร อาจอินทร์, วรพจน์ จักขุพันธ์ และ บุญวัฒน์ สุริยวงศ์, “ระบบสอบย้อนกลับในห่วงโซ่อุปทานข้าวหอมมะลิ โดยการบูรณาการระบบโลจิสติกส์ด้วยเทคโนโลยีสารสนเทศและ RFID,” *การประชุมเชิงวิชาการประจำปีด้านการจัดการโซ่อุปทานและโลจิสติกส์ ครั้งที่ 8 (GTT08)*, 2551, หน้า 35-45.
- [9] วงเดือน พองไธยา, “ความสำคัญของการเพาะเห็ด,” [ออนไลน์] 2556. จาก: <http://www.vcharkarn.com/vblog/38061/1> [5 กุมภาพันธ์ 2556].
- [10] Yu-Chia Hsu, An-Pin Chen and Chun-Hung Wang, “A RFID-Enabled Traceability System for the Supply Chain of Live Fish. *IEEE International Conference on Automation and Logistics (ICAL 2008)*, Qingda, China, 1-3 September 2008, pp.81 - 86.

รูปแบบการรักษาความสัมพันธ์กับพนักงาน บนพื้นฐานความผูกพันของบุคลากร Employee Retention model base on Employee Engagement

ธีรารณ คำวงศ์ (Teeraporn Kamwong)¹ และ ศักดิ์ชัย ตั้งวรรณวิทย์ (Sakchai Tangwannawit)²

^{1,2} ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

Teeraporn.Kam25@gmail.com¹, sakchait@kmutnb.ac.th²

บทคัดย่อ

งานวิจัยเรื่องนี้มีวัตถุประสงค์เพื่อหาแบบการรักษาความสัมพันธ์กับพนักงาน บนพื้นฐานความผูกพันของบุคลากร ภายใต้แนวคิดของการรักษาลูกค้า เครื่องมือที่ใช้ในการวิจัย คือ แบบสอบถามออนไลน์โดยใช้โปรแกรม LimeSurveys และฐานข้อมูล MySQL ในการเก็บข้อมูล แบบสอบถามแบ่งออกเป็น 2 ตอนคือ แบบสอบถามข้อมูลทั่วไป และแบบสอบถามการรักษาความสัมพันธ์กับพนักงาน จัดเก็บข้อมูลจากกลุ่มตัวอย่างจำนวน 200 คน โดยแบ่งข้อคำถามออกเป็น 7 ด้าน ตามทฤษฎี Frederick Herzberg แบบสอบถามเป็นแบบมาตราส่วนประเมินค่า 5 ระดับ กระบวนการวิเคราะห์ข้อมูลแบ่งออกเป็น 2 ส่วนคือ การวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis) เพื่อค้นหาตัวแปรแฝง และการวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Analysis) เพื่อพิสูจน์หรือตรวจสอบสมมุติฐานว่าเป็นไปตามนัยสำคัญทางสถิติหรือไม่

จากการวิเคราะห์องค์ประกอบเชิงสำรวจพบว่า มี 5 องค์ประกอบที่ตัวแปรทุกตัวมีความสัมพันธ์กัน อย่างไรก็ตามด้านความรับผิดชอบในงาน และด้านความน่าสนใจในงาน ยังสามารถแบ่งความสัมพันธ์ได้อีก 2 องค์ประกอบ จากนั้นทำการทดสอบความสอดคล้องของโมเดลในแต่ละองค์ประกอบพบว่า มีค่า P-Value มากกว่า 0.05 ซึ่งโมเดลมีความสอดคล้องกันกับข้อมูลเชิงประจักษ์ และผลการวิเคราะห์องค์ประกอบเชิงยืนยันลำดับที่สอง พบว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์คือ $Chi-Square = 146.87$, $df = 527$, $P-Value = 1.00$, $RMSEA=0.00$

คำสำคัญ: รูปแบบการรักษาความสัมพันธ์, ความผูกพันของบุคลากร

Abstract

The objective of this research was to find the employee retention model base on employee engagement under the concept of customer retention. Online surveys with Limesurveys program and MySQL are employed in this research. The questionnaire was divided into two parts: general information survey and employee retention based on employee engagement. A sample group of 200 people were asked questions relating to the 7 factors of retention developed by Frederick Herzberg theory, the answers are ranked into 5 levels. The process of the analysis is categorized into two parts, which are the Exploratory Factor Analysis to find the latent variable and the Confirmatory Factor Analysis to prove whether or not the hypothesis is statistically significant.

The finding from the Exploratory Factor Analysis indicates that there are 5 factors correlated. However, the job duties and the interest in the job relationships are divided into two components. Afterward, the consistency of the model in each component was tested. The result shows that the P-Value > 0.05, which is consistent with the model with empirical data. Moreover, the model of second-order confirmatory factor analysis proves that the model is also correspond with the empirical data: $Chi-Square = 146.87$, $df = 527$, $P-Value = 1.00$, $RMSEA=0.00$

Keyword: Employee Retention Model, Employee Engagement

1. บทนำ

ความผูกพันของบุคลากรที่มีต่อองค์กรเป็นเรื่องที่ได้รับความน่าสนใจในปัจจุบัน เนื่องจากผลการศึกษายบ่งชี้ว่าพนักงานที่มีความผูกพันจะส่งผลต่อผลดำเนินงานขององค์กร ซึ่งพนักงานที่มีความผูกพันต่อองค์กรจะทำงานหนักขึ้น เพิ่มระดับความพยายามทุ่มเท ถ้าองค์กรไม่สามารถรักษานักงานและตอบสนองความต้องการที่หลากหลายได้ ก็จะส่งผลกระทบต่อความผูกพัน แรงจูงใจในการปฏิบัติงาน และนำไปสู่การลาออกในที่สุด ปัญหาที่เกิดขึ้นคือ องค์กรยังไม่มีการทำในเรื่องของความผูกพันที่ชัดเจน ส่วนใหญ่จะมุ่งเน้นไปทางด้านความพึงพอใจของพนักงาน ทำให้การบริหารจัดการงานทางด้านทรัพยากรบุคคลยังไม่ประสบผลสำเร็จ และการบริหารจัดการกับบุคลากรในองค์กรเรื่องของการสร้างความผูกพันและการรักษานักงานยังไม่มีรูปแบบการดำเนินงานที่ชัดเจน

จากปัญหาดังกล่าวข้างต้นทำให้ผู้วิจัยเล็งเห็นถึงความสำคัญในการศึกษาหารูปแบบของการรักษาความสัมพันธ์ บนพื้นฐานของความผูกพันของบุคลากร โดยแบ่งปัจจัยออกเป็น 7 ด้านตามทฤษฎีการจูงใจและรางวัลรักษาของ Herzberg ภายใต้แนวคิดของการรักษาลูกค้ามาประยุกต์ใช้ในงานวิจัย ขั้นตอนของการวิเคราะห์ได้นำเทคนิคของการวิเคราะห์องค์ประกอบเชิงสำรวจ และการวิเคราะห์องค์ประกอบเชิงยืนยันมาใช้ในการวิเคราะห์ปัจจัย เพื่อสร้างและทดสอบความสอดคล้องของโมเดล

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1. ความหมายของความผูกพันของพนักงาน

ความผูกพันของพนักงานในองค์กรหมายถึงความมุ่งมั่นและความสามารถที่จะอุทิศตนเพื่อความสำเร็จขององค์กร อุทิศเวลา สติปัญญา และแรงงาน ความรู้สึกถึงเป้าหมาย หวังและภูมิใจ ซึ่งทำให้พยายามอย่างสุดความสามารถในการทำงาน [1]

2.2. ความหมายของทฤษฎีการจูงใจและรางวัลรักษา

Frederick Herzberg [2] ได้อธิบายว่ามีปัจจัย 2 ปัจจัยที่มีผลต่อการจูงใจพนักงานให้ทำงานคือ 1) Hygiene Factors เป็นปัจจัยที่ต้องมีในการทำงาน ถ้าขาดส่วนนี้ไปจะมีปัญหาเกี่ยวกับพนักงาน แต่ปัจจัยนี้ไม่ได้เป็นปัจจัยที่จูงใจให้พนักงานอยาก

ทำงานโดยตรง และ 2) Motivator Factors เป็นปัจจัยที่สามารถจูงใจพนักงานให้อยากทำงาน โดยแบ่งออกเป็น ด้านการบังคับบัญชา ด้านความรับผิดชอบในงานที่ทำ ท้าย ด้านความสำเร็จในการทำงาน ด้านความน่าสนใจในงาน ด้านความก้าวหน้าในงาน ด้านการได้รับความสำคัญ และด้านเงินเดือนและสวัสดิการ

2.3. วงจรชีวิตการเป็นลูกค้า (Customer Life Cycle)

วงจรชีวิตการเป็นลูกค้าเป็นกระบวนการในการบริหารความสัมพันธ์กับลูกค้าอย่างต่อเนื่อง ตั้งแต่เริ่มมาเป็นลูกค้าขององค์กร การมอบคุณค่าของผลิตภัณฑ์หรือบริการเพื่อตอบสนองความต้องการของลูกค้า ไปจนถึงการรักษาลูกค้าไว้ให้อยู่กับองค์กร ซึ่งประกอบด้วยปัจจัยหลัก 3 ประการ 1) การหาลูกค้าใหม่เข้าองค์กร (Customer Acquisition) โดยการสร้างความเด่นของสินค้าหรือบริการที่ใหม่และเสนอความความสะดวกสบายให้กับลูกค้า 2) การรักษาลูกค้า (Customer Retention) ให้อยู่กับองค์กรนานที่สุด และการดึงลูกค้าให้กลับมา ใช้สินค้าหรือบริการ โดยฟังความคิดเห็นจากลูกค้าและพนักงานในองค์กร รวมถึงการเสนอสินค้าและบริการใหม่ 3) การเพิ่มความพึงพอใจให้กับลูกค้าเพื่อทำการซื้อสินค้าและบริการ โดยผ่านขั้นตอนการทำงานที่กระชับเพื่อการสนองความต้องการของลูกค้าได้รวดเร็วและถูกต้อง และตอบสนองสิ่งที่ลูกค้าต้องการหรือเสนอสิ่งที่ลูกค้าต้องการ ผ่านหน่วยงาน ลูกค้าสัมพันธ์ (Customer Service) [3]

2.4. ความหมายของการวิเคราะห์องค์ประกอบ

การวิเคราะห์องค์ประกอบ (Factor Analysis) หรือบางครั้งเรียกว่า การวิเคราะห์ปัจจัยเป็นเทคนิคที่จะจับกลุ่มหรือรวมตัวแปรที่มีความสัมพันธ์กันไว้ในกลุ่ม หรือปัจจัยเดียวกัน ตัวแปรที่อยู่ในปัจจัยเดียวกันจะมีความสัมพันธ์กันมาก โดยความสัมพันธ์นั้นอาจเป็นไปในทิศทางบวก หรือทิศทางลบ ส่วนตัวแปรที่อยู่คนละปัจจัยจะไม่มีความสัมพันธ์กัน [4]

ภาพที่ 1: Exploratory Factor Analysis : EFA

รูปภาพที่ 1 แสดงให้เห็นรูปแบบการวิเคราะห์องค์ประกอบเชิงสำรวจ ในรูปนี้ตัวแปรในสี่เหลี่ยมคือตัวแปรที่สังเกตได้และตัวแปรในวงกลมคือตัวแปรแฝงหรือองค์ประกอบ ถูกสรุตรงซึ่งจากตัวแปรแฝงไปยังตัวแปรที่สังเกตได้ แสดงให้เห็นความสัมพันธ์ของตัวแปรแฝงที่ส่งผลต่อตัวแปรที่สังเกตได้ เส้นโค้งระหว่างตัวแปรแฝง 2 ตัวแสดงให้เห็นถึงความสัมพันธ์ระหว่างตัวแปรแฝง รูปแบบกระบวนการวิเคราะห์เชิงสำรวจ สามารถแบ่งได้ 4 ขั้นตอนคือ 1) เก็บข้อมูลและสร้างเมตริกสหสัมพันธ์ 2) เลือกวิธีการหมุนแกน 3) เลือกค่า loading 4) ตั้งชื่อองค์ประกอบที่วิเคราะห์ได้ [4]

2.5. วัตถุประสงค์ของการวิเคราะห์องค์ประกอบ

2.5.1. เพื่อสำรวจหรือค้นหาตัวแปรแฝง ที่ซ่อนอยู่ภายใต้ตัวแปรที่สังเกตหรือวัดได้เรียกว่าการวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis : EFA)

2.5.2. เพื่อพิสูจน์ตรวจสอบหรือยืนยันทฤษฎีที่ผู้อื่นค้นพบเรียกว่าการวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Analysis : CFA)

ภาพที่ 2: Confirmatory Factor Analysis : CFA

2.6. การสร้างโมเดลการวัดองค์ประกอบ

การสร้างและทดสอบ โมเดลการวัดองค์ประกอบ เป็นขั้นตอนของการวิเคราะห์องค์ประกอบเชิงยืนยัน ต้องเริ่มต้นที่สร้างเมตริกสหสัมพันธ์ หรือเมตริกความแปรปรวนความแปรปรวนร่วม หรือเมตริกอื่น ๆ ที่มีลักษณะคล้ายกัน ส่วนจุดประสงค์ในการทดสอบโมเดล ซึ่งจะต้องตั้งอยู่บนพื้นฐานของทฤษฎีหรือข้อมูลที่มีอยู่ สมมติฐานจะต้องตั้งให้เหมาะสมกับข้อมูล โมเดลจะต้องกำหนดระดับของความสัมพันธ์ระหว่างคู่ของตัวแปรแต่ละตัวกับองค์ประกอบ ความแตกต่างของ

พารามิเตอร์คงที่และอิสระ ในการวิเคราะห์องค์ประกอบเชิงยืนยันสามารถกระทำได้ [5] โดยใช้โปรแกรมคอมพิวเตอร์ที่หลากหลาย เช่น LISREL for Windows, Amos เป็นต้น

การทดสอบโมเดล จะถูกทดสอบด้วยโปรแกรมคอมพิวเตอร์ ในเรื่องของความสมบูรณ์ในการวิเคราะห์นั้นผลลัพธ์ที่ได้จะต้องแสดงค่าสถิติที่แตกต่างกันหลายค่า สำหรับใช้ในการทดสอบความสอดคล้องของโมเดล หรืออธิบายความแปรปรวนร่วมระหว่างตัวแปร ค่าสถิติจะถูกนำเสนอในรูปของ “Fit statistics” ซึ่งค่าทั้งหมดจะถูกประมวลและแสดงออกมาในคราวเดียวกัน สถิติเหล่านี้จะใช้ในการประเมินความสอดคล้องของโมเดลกับข้อมูลเชิงประจักษ์ และใช้อธิบายความสัมพันธ์ระหว่างตัวแปรที่สังเกตได้กับตัวแปรแฝง

2.7. โมเดลลิสเรล

โมเดลลิสเรลเป็น โมเดลที่แสดงความสัมพันธ์โครงสร้างระหว่างตัวแปรทั้งหมดมาจากคำภาษาอังกฤษว่า “ Linear Structure RELationship : LISREL model ” นอกจากนี้โมเดลลิสเรลยังมีชื่อว่า “ โมเดลสมการโครงสร้างเชิงเส้น ” (linear structure equation model) และ “ โมเดลโครงสร้างความแปรปรวนร่วม ” (covariance structure model) ผู้ที่พัฒนาโปรแกรมลิสเรล (Linear Structure RELationship : LISREL model) คือ K.G. Joreskog และ D. Sorbom ในช่วงปีค.ศ.1967 – 1979 โปรแกรมลิสเรลนี้ นับเป็นโปรแกรมแรกที่ได้รับการพัฒนาขึ้นเพื่อวิเคราะห์โมเดลลิสเรล

2.8. ลักษณะการวิเคราะห์ด้วยโปรแกรมลิสเรล

การวิเคราะห์ด้วยโปรแกรมลิสเรลเป็นการวิเคราะห์ข้อมูลสำหรับการวิจัยที่มีโมเดลเชิงสาเหตุ (causal-model) มีตัวแปรแฝง (latent variables) ที่ตัวบ่งชี้หลายตัวมีการประมาณค่าความคลาดเคลื่อนได้ (ทำให้ผ่อนคลายข้อตกลงเบื้องต้นของโมเดลการวิเคราะห์สาเหตุแบบเดิม) และความสัมพันธ์ระหว่างเศษเหลือ ผลการวิเคราะห์มีความถูกต้องแม่นยำ อีกทั้งโมเดลลิสเรลยังมีคุณลักษณะทั่วไป ครอบคลุมวิธีการวิเคราะห์ข้อมูลทางสถิติขั้นสูงเกือบทุกรูปแบบ ทั้งการวิเคราะห์ความแปรปรวน (analysis of variance) หรือ ANOVA การวิเคราะห์การถดถอยพหุคูณ (multiple regression analysis) การวิเคราะห์ความแปรปรวนร่วม (analysis of covariance) สถิติวิเคราะห์ตัวแปรพหุนาม (multi-sample analysis or multi group analysis)

การวิเคราะห์องค์ประกอบ(factor analysis) การวิเคราะห์อิทธิพล (path analysis) การวิเคราะห์สหสัมพันธ์คาโนนิกอล (canonical correlation analysis) นักวิจัยที่มีความรู้ความเข้าใจเทคนิคการวิเคราะห์โมเดล ลิสเรลจะเข้าใจลักษณะร่วมกันของสถิติวิเคราะห์ดังกล่าว และใช้ลิสเรลวิเคราะห์ข้อมูลได้ทั้งหมด

2.9. ผลงานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องกับการสำรวจความคิดเห็นปัจจัยที่มีผลต่อการรักษาพนักงาน พบว่าพนักงานให้ความสำคัญกับปัจจัย ด้านความมั่นคงปลอดภัย ด้านค่าตอบแทนและสวัสดิการ ด้านความยุติธรรมในการเลื่อนตำแหน่ง และปัจจัยด้านความผูกพันต่อองค์กร ซึ่งส่งผลกระทบต่อการจูงใจ ความผูกพันและการรักษาพนักงาน [6, 7, 8]

งานวิจัยที่เกี่ยวข้องกับการวิเคราะห์ปัจจัย โดยใช้วิธีการวิเคราะห์องค์ประกอบเชิงสำรวจ และการวิเคราะห์องค์ประกอบเชิงยืนยันชั้นที่สอง เพื่อสังเคราะห์รูปแบบโมเดลองค์ประกอบของเกณฑ์สมรรถนะที่มีความสำคัญต่อนักเทคโนโลยีสารสนเทศ [9]

ดังนั้นผู้วิจัยจึงเล็งเห็นว่าการนำการวิเคราะห์ปัจจัยมาวิเคราะห์หาองค์ประกอบด้านการรักษาความสัมพันธ์กับด้านความผูกพันของบุคลากร เพื่อให้องค์กรรับทราบถึงปัจจัยที่สำคัญต่อการรักษาพนักงานและปัจจัยด้านความผูกพัน

3. วิธีดำเนินการวิจัย

การดำเนินงานวิจัยนี้เป็นการวิเคราะห์หารูปแบบของการรักษาความสัมพันธ์กับพนักงานบนพื้นฐานความผูกพันของบุคลากร โดยใช้วิธีการวิเคราะห์องค์ประกอบเชิงสำรวจ และการวิเคราะห์องค์ประกอบเชิงยืนยัน ผู้วิจัยได้กำหนดขั้นตอนของการดำเนินงาน ออกเป็นรายละเอียดดังนี้

3.1. ศึกษาข้อมูลและสร้างเครื่องมือที่ใช้ในการวิจัย

ผู้วิจัยได้ทำการศึกษาทฤษฎีที่เกี่ยวข้องกับการรักษาความสัมพันธ์ และความผูกพันของบุคลากร จากงานวิจัย บทความ วารสาร ที่เกี่ยวข้อง และได้ทำการสร้างแบบสอบถาม โดยแบ่งปัจจัยหลักของข้อคำถามออกเป็น 7 ด้านตามทฤษฎีการจูงใจและธำรงรักษา ภายในปัจจัยหลักจะประกอบไปด้วยข้อคำถามที่เกี่ยวข้องกับการรักษาความสัมพันธ์ และความผูกพันไว้ด้วยกัน รวมข้อคำถาม

ทั้งหมด 34 ข้อ นำข้อคำถามไปหาคุณภาพของแบบสอบถาม โดยความคิดเห็นจากผู้เชี่ยวชาญทั้งหมด 5 ท่าน นำมาคำนวณค่า IOC ซึ่งนำข้อคำถามที่ใช้ได้ทั้งหมด ไปจัดทำเป็นแบบสอบถามออนไลน์



ภาพที่ 3: แบบสอบถามออนไลน์ด้วยโปรแกรม LimeSurveys

3.2. วิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis : EFA)

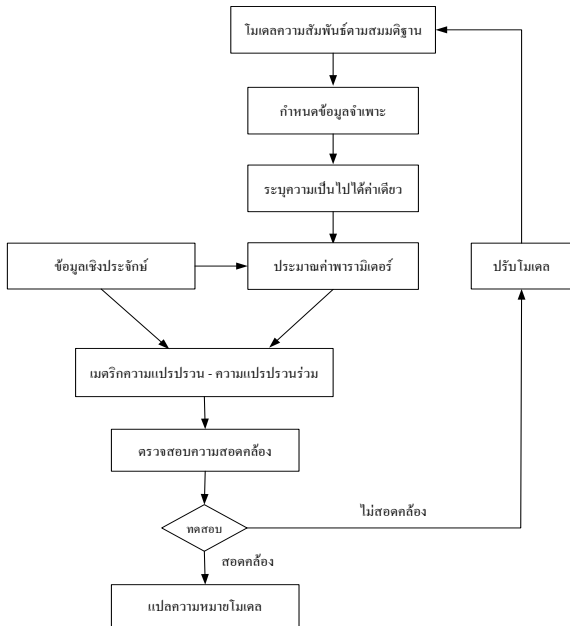
นำข้อมูลทั้งหมดจากฐานข้อมูลที่ได้จากการตอบแบบสอบถามมาทำการวิเคราะห์เพื่อสกัดปัจจัยด้วยโปรแกรม SPSS ใช้วิธี Principal Components Analysis (PCA) และวัดค่า Kaiser-Meyer-Olkin (KMO) ค่าที่ควรจะต้องมีค่า > 0.5 เข้าใกล้ 1 ทำให้การวิเคราะห์ปัจจัยเหมาะสมกับข้อมูลในแต่ละด้าน และหาค่าความแปรปรวนของตัวแปร ค่าไอเกน (Eigen value) จะต้องมีค่า > 1 เลือกวิธีการหมุนแกนแบบ Varimax

3.3. วิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Analysis : CFA)

นำค่าสถิติของแต่ละปัจจัยที่ได้จากการวิเคราะห์องค์ประกอบเชิงสำรวจมาสร้างเป็นโมเดล ด้วยโปรแกรม Lisrel เพื่อทดสอบความสอดคล้องของโมเดล ว่าโมเดลตามทฤษฎีสอดคล้องกับข้อมูลเชิงประจักษ์ แบ่งการวิเคราะห์ออกเป็น 2 ส่วนคือ 1) การวิเคราะห์ความสอดคล้องของโมเดลในแต่ละปัจจัย และ 2) การวิเคราะห์องค์ประกอบเชิงยืนยันลำดับที่สอง เริ่มจากนำข้อมูลที่ได้จากการวิเคราะห์มาสร้างเมตริกซ์ความสัมพันธ์ ระหว่างคู่ของตัวแปรทุกตัว (Correlation Matrix) เพื่อสร้างเป็นโมเดลขององค์ประกอบในแต่ละด้าน และนำปัจจัยทั้งหมดมาวิเคราะห์โมเดลเชิงยืนยันลำดับที่สอง

3.4. ทดสอบความสอดคล้องของโมเดล

ทำการทดสอบความสอดคล้องของโมเดลการวัด ขององค์ประกอบด้านต่างๆ เพื่อดูโมเดลตามทฤษฎีที่ได้จากการวิเคราะห์องค์ประกอบเชิงสำรวจ สอดคล้องกับข้อมูลเชิงประจักษ์ที่ได้จากการเก็บรวบรวมข้อมูลจากกลุ่มตัวอย่างหรือไม่ ซึ่งถ้าไม่สอดคล้องก็สามารถที่จะปรับโมเดลช่วยจนสอดคล้อง



ภาพที่ 4: การทดสอบความสอดคล้องของโมเดล

4. ผลการวิจัย

ผลการดำเนินงานวิจัยโดยวิธีการวิเคราะห์องค์ประกอบเชิงสำรวจ และการวิเคราะห์องค์ประกอบเชิงยืนยัน ทั้งหมด 7 ด้าน ดังนี้

ตารางที่ 1: ค่าของการวิเคราะห์องค์ประกอบเชิงสำรวจ และการวิเคราะห์องค์ประกอบเชิงยืนยัน

Factors	EFA	CFA			
	Eigenvalue > 1	Chi-square	df	P-Value	RMSEA
ด้านการบังคับบัญชา	1	13.99	8	0.082	0.061
ด้านความรับผิดชอบ	2	3.64	3	0.302	0.033
ด้านความสำเร็จ	1	4.36	2	0.112	0.077
ด้านความน่าสนใจ	2	2.34	2	0.310	0.029
ด้านความก้าวหน้า	1	0	0	1.000	0.000
ด้านการได้รับความสำคัญ	1	1.80	1	0.179	0.064
ด้านเงินเดือนและสวัสดิการ	1	6.46	3	0.091	0.076

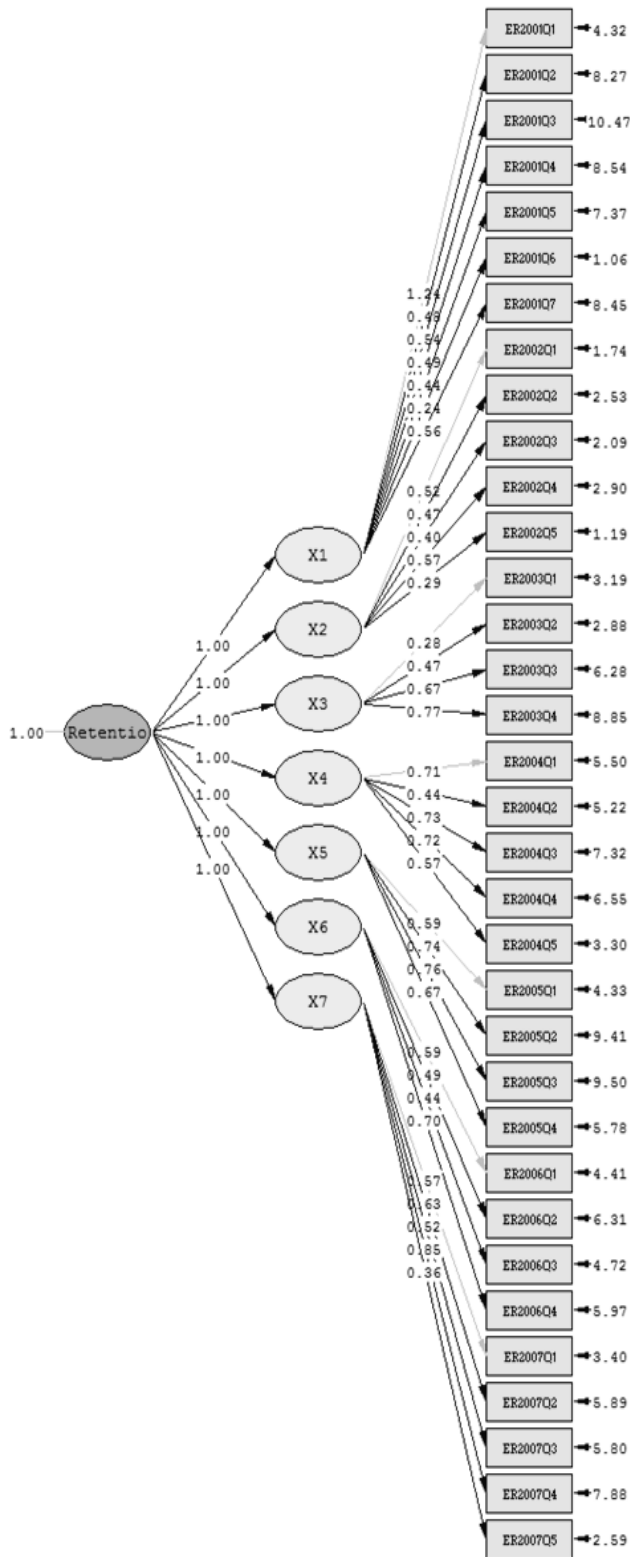
ผลการทดสอบการวิเคราะห์องค์ประกอบเชิงสำรวจพบว่า ด้านความรับผิดชอบในงาน และด้านความน่าสนใจในงานที่มี

ค่า eigenvalue >1 มี 2 องค์ประกอบ แปลผลได้ว่าด้านความรับผิดชอบในงาน สามารถแบ่งกลุ่มได้ 2 ด้านคือ การทุ่มเทให้กับองค์กร และสิ่งที่องค์กรมอบให้ ส่วนด้านความน่าสนใจในงาน สามารถแบ่งกลุ่มได้ 2 ด้านคือ ความเหมาะสมในงานที่ทำ และความสอดคล้องของงานและชีวิตส่วนตัว

จากข้อมูลข้างต้นนำมาวิเคราะห์องค์ประกอบเชิงยืนยันในแต่ละปัจจัย เพื่อพิจารณาจากความสอดคล้องทางทฤษฎี พบว่าค่า P-Value > 0.05 แสดงว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์ ดังนั้นจึงนำมาทำการวิเคราะห์องค์ประกอบเชิงยืนยันลำดับที่สองจากปัจจัยทั้งหมด 7 ด้าน ได้ค่าไค-สแควร์เท่ากับ 146.87 (Chi-Square) ที่องศาอิสระ(df) เท่ากับ 527 และค่า P เท่ากับ 1.00 ค่าดัชนีวัดระดับความสอดคล้อง(GFI) เท่ากับ 0.96 และการวิเคราะห์องค์ประกอบเชิงยืนยันลำดับที่สอง ได้ค่าดัชนีวัดระดับความสอดคล้องที่ปรับแก้ (AGFI) เท่ากับ 0.95 ผลการปรับแก้โมเดลพบว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์อย่างสมบูรณ์ และได้รูปแบบในการรักษาความสัมพันธ์กับพนักงานบนพื้นฐานความผูกพันของบุคลากร แสดงดังภาพที่ 5

5. สรุปและอภิปรายผล

จากผลการวิจัยได้รูปแบบของการรักษาความสัมพันธ์ในแต่ละปัจจัยดังนี้ 1) ด้านการบังคับบัญชา พนักงานพึงพอใจในผู้บังคับบัญชาของตนเอง แต่ควรแก้ไขในเรื่องการติดตามงานของผู้ใต้บังคับบัญชาอย่างใกล้ชิด 2) ด้านความรับผิดชอบในงาน พนักงานมีโอกาสใช้ความรู้ความสามารถอย่างเต็มที่ แต่ยังคงขาดโอกาสการตัดสินใจในงานที่น้อย 3) ด้านความสำเร็จในการทำงาน พนักงานมีโอกาสได้เรียนรู้และมีความเติบโตในหน้าที่ แต่ผลความสำเร็จของงานยังไม่ชัดเจน 4) ด้านความน่าสนใจในงาน พนักงานมีโอกาสได้ใช้ความรู้ ความสามารถ และความคิดริเริ่มสร้างสรรค์ แต่บางงานยังไม่สอดคล้องกับนโยบาย 5) ด้านความก้าวหน้าในงาน การพิจารณาเลื่อนตำแหน่งมีความเหมาะสม แต่ยังคงขาดความยืดหยุ่นและการมีส่วนร่วมในการตัดสินใจในงาน 6) ด้านการได้รับความสำคัญ พนักงานได้รับคำชื่นชมจากเพื่อนร่วมงาน แต่ได้รับการชื่นชมจากลูกค้าน้อย และ 7) ด้านเงินเดือนและสวัสดิการ พนักงานได้รับการดูแลเรื่องสุขภาพเป็นอย่างดี ควรปรับปรุงในส่วนของบริษัทและสภาพการทำงานให้ดีขึ้น



ภาพที่ 5: ผลการวิเคราะห์องค์ประกอบเชิงยืนยันลำดับที่สอง

6. เอกสารอ้างอิง

- [1] อภิชัย ศรีเมือง, “ความหมายของความผูกพันของพนักงาน”, *วารสารการบริหารสำหรับนักบริหารงานบุคคลมืออาชีพ ปีที่ 26* สมาคมการจัดการงานบุคคล, 2549.
- [2] F. Herzberg, “The Motivation to Work”, *New York: John Wiley and Sons*, 1959.
- [3] กัทร รุจยาชยะกูร, “เจาะลึกการสร้างฐานลูกค้าใหม่ใน Customer Lifecycle”, สืบค้นวันที่ 15 มกราคม 2557 จาก <http://www.incquity.com/articles/customer-lifecycle-management/>
- [4] รสริน ศรีริگانนท์, “การวิเคราะห์องค์ประกอบ(Factor Analysis)”, *เอกสารประกอบการบรรยาย มหาวิทยาลัยราชภัฏสงขลา*, 2555.
- [5] ฉัตรศิริ ปิยะพิมลสิทธิ์, “การวิเคราะห์องค์ประกอบเชิงยืนยัน”, *วารสารการวัดผลการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ*, 2554.
- [6] กนกพร สุโพภาค, “การชำระรักษาพนักงานของกลุ่มตัวอย่างพนักงานที่ทำงานในเขตกรุงเทพมหานคร”, *มหาวิทยาลัยบูรพา*, 2553.
- [7] ชันย์ชนก ศรีสวัสดิ์, “การศึกษาแนวทางปฏิบัติในการชำระรักษาบุคลากรที่มีสมรรถนะสูง: กรณีศึกษาสำนักงานคณะกรรมการข้าราชการพลเรือน”, *วารสารวิทยบริการ ปีที่ 24 ฉบับที่ 3 หน้า 47-58*, 2556.
- [8] เสกสรร ศิริวรรณ, “ปัจจัยที่มีผลต่อความผูกพันต่อองค์กรของพนักงานบริษัทซีพีเอฟ เทคดิ่ง จำกัดในเขตภาคเหนือ”, *วิทยานิพนธ์ บริหารธุรกิจมหาบัณฑิต วิทยาลัยพาณิชยศาสตร์ มหาวิทยาลัยบูรพา*, 2554.
- [9] มณเชิธร รัตนศิริวงศ์วุฒ และทีมงาน, “ปัจจัยที่มีอิทธิพลต่อสมรรถนะนักเทคโนโลยีสารสนเทศโดยการวิเคราะห์องค์ประกอบเชิงยืนยันลำดับที่สอง”, *ปัญหาพิเศษ วิทยาศาสตร์มหาบัณฑิต คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*, 2553.

AUTHOR INDEX

	Page
Adisak Sangsongfa	450
Aekkarat Lorphichian	223
Aiyapan Eagobon	1012
Akarapol Padungdit	625
Amorntape Yamkasorn	45
Anantaporn Hunskunatai	95
Angkana Suwanjatuporn	305
Anukul Butprom	176
Anupan Rodtook	8
Anupap Meesomboon	111
Anusorn Untao	110
Anuwat Jaidee	193
Aongart Aun-A-Nan	252
Apichat Lagdee	352,914
Apichaya Kiartubolpaiboon	310
Apiwat Sawatdirat	792
Assadaporn Sapsomboon	846
Athaporn Gingboo	13
Athita Onuean	865
Atiporn Panpum	425
Attapol Kunlerd	7
Aumaporn Somthong	595
Aumpai Chobdean	64
Autthapon Hutthapanit	859
Baramee Osateerakul	462
Benjaporn Paripinidchai	1006
Benjawan Intara	677
Boonhatai Kruekaew	694
Boonserm Kijisirikul	466
Burin Rujjanapan	299
Chaiyaporn Khemapatapan	192, 204, 504
Chamaiporn Tongkaw	774
Chanadej Yamcee	204
Chanokthamon Suksri	631
Chanyaphon Prasomsat	504
Chaowalit Somboonpattanakit	240
Chaowat Prattanasangkusol	896
Charinthorn Aumgril	952
Charun Sanrch	698
Chatchawin Namman	246

AUTHOR INDEX

	Page
Chate Patanothai	713, 810, 908
Chayaphol Prapaipornlert	683
Chonnikarn Rodmorn	731
Choopan Rattanapoka	109
Chutima Ampairat	229
Dechawut Wanichsan	780
Duangjai Noolek	822
Duangnet Yotkamlue	378
Duangporn Jeamumporn	737
Dusanee Supawantanakul	358, 372
Ekapong Chuasuwan	64
Goragod Pangsomboon	316
Guangqian Chen	13
Hathai Kaewkorn	613
Hatsada Phruekthachanbodi	217
Isaraporn Vithyaviranont	478
Isranol Polparee	834
Jakkarin Suksawatchon	515, 755
Jakkree Srinonchat	792
Janasthar Srisuan	95
Janejira Wanglee	156
Jaroonluk Suporn	407
Jatupit Aiyarakom	76
Jaturada Deeying	846
Jirapong Pongpilasarn	985
Jitsaran Seekuka	601
Jonathan Chan	583
Julaluk Watthananon	975
Jutarat Tangkittiwat	26
Kanda Rongsawad	547
Kanisorn Jeekratok	927
Kannika Nutchomphu	671
Kanokwan Atchariyachanvanich	840
Khamron Sunat	407, 595, 969
Khemika Moosikaphan	694
Khwanchai Suksaen	276
Khwanruan So-Ubol	150
Kiattisak Yochanang	559
Kitsuchart Pasupa	115, 258, 683
Kittaboon Panjarattankorn	115

AUTHOR INDEX

	Page
Kitti Surpare	322
Kittichai Sopitwetmontree	83
Kittinan Noimanee	146
Kittipun Khantirirat	683
Kittisak Onuean	762
Kittisak Thongyuan	798
Krit Sirisup	884
Krittachai Boonsivanon	553, 577
Krittika Tipyodsree	497
Ladda Preechaveerakul	156
Laongtip Mathurasa	358, 372
Lawankorn Soimart	725
Mahasak Ketcham	462, 798, 1006
Maleerat Sodanil	70, 168, 293, 334, 559, 658 671, 737, 774, 896
Manassanant Chantho	276
Maneevan Supatimassaro	840
Mathawee Sutikun	109
Mathuros Panmuang	731
Matupoom Kittisuntaropast	652
Mayuree Jeeramard	270
Meechai Lohakan	322
Metawee Jindarak	828
Metha Sirigool	210
Monchai Tiantong	39, 358, 372
Mongkol Thongkraikaew	589
Monsinee Attasiripanya	472
Montean Rattanasiriwongwut	877, 982, 994
Nadh Ditcharoen	419
Nalinpat Porrawatpreyakorn	26, 877
Namthip Marknakorn	168
Narakorn Srivapee	431
Narathip Thongpan	484
Narongdech Keeratipranon	638
Narumol Fengseekom	665
Narumol Nuanchawee	920
Natawut Nupairoj	706
Natchasit Chukiatkhajorn	146
Nattaphon Saelim	908
Nattavadee Hongboonmee	287, 940
Nattavee Utakrit	619, 884

AUTHOR INDEX

	Page
Nattawut Moonsart	276
Natthawan Phoonson	859
Natthawoot Punroob	13
Natwalai Komkham	535
Nawaporn Wisitpongphan	484, 528
Netchanok Jansawang	174
Ngamnij Arch-Int	127
Nikhom Wuttiya	969
Nilubon Tongchai	444
Nipawan Sapet	652
Nitikorn Intarasod	994
Nitipat Petmongkonjaras	522
Nongluk Covavisaruch	920
Nonlapan Praloppan	419
Noppamas Pukkhem	749
Nuengwong Tuaycharoen	162, 316, 1012
Nunnapus Benjamas	652, 665
Nuntanee Chuaychoo	58
Nutchuda Mongkolchart	528
Nutkamol Sripum	396
Nutthavipa Songsook	20
Nutthreeya Songyai	413
Paiboon Chatthong	559
Paisan Simalaotao	957
Pamok Rattanatriyapibal	70
Panich Sudkhot	677
Panida Tunyanon	816
Panwit Tuwanuti	198, 210, 396, 625
Paranee Apiromsanee	305
Pariwat Kumkunkum	859
Parkpoom Pornpratanwech	198
Patchai Phuangsuwan	491
Patharawut Saengsiri	631
Patsapol Prommat	902
Peerachai Chinporn	346
Peerapad Pandison	64
Phanarut Srichetta	645, 768
Phayung Meesad	450, 509, 528, 541, 762, 804
Phichitra Siriwat	940
Phongkorn Aksonnittrakul	328

AUTHOR INDEX

	Page
Phonpimon Panthuvapee	768
Photjane Sukchaona	39
Phumin Kirawanich	1018
Pichayotai Mahathanapawat	89
Piradee Chalomprattanapong	871
Pitipark Pinrod	390
Piyawan Sanpanya	595
Pongpat Singsri	515
Pongpisit Wuttidittachotti	694, 719, 871
Pornchai Mongkolnam	583
Pornnapa James	631
Pornsiri Muenchaisri	497, 853
Prachyanun Nilsook	390
Pradit Pitaksathienkul	235
Pragaimas Srisuktaksin	281
Prapada Pokumapichai	982
Prawit Oungcharoon	316
Prudtipong Pengsiri	528
Pudsadee Boonrawd	281, 456, 535, 834, 890, 902
Pusadee Seresangtakul	7, 1024
Rakwarinn Wannasin	553
Rapeepong Rattanawaorahirunkul	601
Rapheephong Sriwiboon	810
Rattana Wetprasit	340
Ratchaporn Suphap	1000
Ratthasilp Ranokphanuwat	316
Rungrat Wiangsripanawan	589
Rushapon Piaboran	192
Saiyan Saiyod	127
Sakchai Tangwannawit	270, 328, 988, 1000, 1030
Saksit Sawang	396
Sangavood Pota	246
Sapa Chanyachatchawan	316
Sarach Sriswadpong	180
Sarawit Timim	264
Sarawut Supannarach	601
Sarawuth Rungcharoenkit	258
Sarayut Neankratok	401
Sarita Khuptayanon	133
Sathit Suwannawach	780

AUTHOR INDEX

	Page
Sattarpoom Thaiparnit	462
Siitiporn Ittasombut	365
Sinate Kimseng	607
Sippawit Suggree	625
Sirapat Boonkrong	217,223 , 229, 235, 240
Sirapat Chiewchanwattana	407,595 , 969
Sirikorn Krompho	877
Siripong Matame	264
Siriruang Patchuay	484
Sirirut Vanichayobon	607,613
Sitthipon Limchaichada	853
Somjate Srinakorn	658
Somjet Poosri	352
Somjit Arch-Int	127
Sompong Valuvanathon	571
Somrudee Jaroenwaree	293
Somrudee Kaewkao	140
Sorakis Srikasem	786
Suchada Ratanakongnate	83
Sukree Sinthupinyo	1
Sumalee Eisariyodom	32
Sumet Darapisut	755
Sumet Prabhavat	180,186
Sumet Siriruangampai	706
Sunantha Sodsee	252,725
Sunisa Rimcharoen	934
Sunti Preechawit	1018
Suntis Jullapho	619
Sununta Klintawon	174
Supachoke Pongquan	186
Supakit Nootyaskool	472,828
Supaporn Kansomkeat	58
Supaporn Kurdkit	725
Supaporn Porntra	419
Supitchaya Janplong	840
Supot Nitsuwat	737
Supote Bualerng	101
Surachat Boonsaen	162
Suraphong Cheawsakunwattana	1
Surapon Kraijak	210

AUTHOR INDEX

	Page
Surat Duangchatom	384,431
Sureewan Jangjit	484
Suriya Pinitkan	528
Sutheera Phannack	334
Suwat Pinyopan	466
Tanasanee Phienthrakul	522
Tanchanok Toopkird	665
Tanong Chuvanitchanon	180
Tanyada Taychatanont	652
Tanyamai Puchanapakorn	456
Tawa Khampachua	689
Tawatchai Ruangtanaturak	127
Taweesak Noipasee	689
Teerapong Sungsi	963
Teeraporn Kamwong	1030
Thammarsat Visutarom	583
Thanapon Boonsanpol	638
Thanate Yuensook	437
Thaphat Mupattararot	121
Tharach Arreerard	372,437
Thararat Udomchaibanjerd	890
Thawatchai Sahapong	437
Theerapong Boriboon	491
Therdpong Daengsi	694,719
Thippaya Chintakovid	822
Thiraphat Pholsri	645
Thitikorn Prakrongyat	419
Thodsaphon Prachpreecha	786
Titapat Utthakan	743
Tuul Triyason	719
Twittie Senivongse	51, 76, 140
Ubolwan Chaovanakij	115
Ubonrat Sirisukpoca	957
Upady Hatthasin	264
Ureerat Suksawatchon	32, 515
Usa Choothong	287
Usanad Ua-Apisitwong	401
Veera Muangsin	121,478
Veeraporn Siddoo	340
Veerawan Janthanasub	804

AUTHOR INDEX

	Page
Vishit Shuuttha	645
Vitthawat Sajjacharoenpong	51
Walongkorn Pasakanon	565
Wannasiri Thurachon	20
Warangkhan Kimpan	700
Waranyu Wongseree	310
Wararat Songpan	101
Watchara Chatwiriya	547
Watcharapat Laonawatthana	316
Werayut Pimpaporn	541
Winyoo Uttara	384
Wipawan Buathong	509
Wirinya Songcheang	287
Wisinee Prasiditthaikul	472
Wongkot Sriurai	413,425
Worachat Sukra	749
Woramet Withunwet	346
Worapapha Arreerard	174, 346, 352, 358, 365, 378, 384, 431, 437, 946
Worapong Ratanaprasit	713
Worasit Choochaiwattana	150, 816
Worawat Akapan	719
Worawut Yimyam	252
Worraphot Sumethawatthanaphong	20
Worrawat Narongkhachavana	180,186
Yachai Limpiyakorn	133
Zagon Budsabong	89



THE 11TH NATIONAL CONFERENCE ON
COMPUTER AND INFORMATION
TECHNOLOGY

7TH - 8TH MAY 2015

FACULTY OF INFORMATION TECHNOLOGY
KING MONGKUT'S UNIVERSITY OF TECHNOLOGY NORTH BANGKOK

WWW.NCCIT.NET

CONFERENCE ORGANIZERS



ECU - Australia



OSU - USA



Fern - Germany



KMUTNB - Thailand



HNUE - Vietnam



Chemnitz - Germany



UBU - Thailand



RMU - Thailand



KRU - Thailand



NPRU - Thailand



MSU - Thailand



RMUTL - Thailand



RMUTK - Thailand



RMUTT - Thailand



PSU - Thailand



NIDA - Thailand

KING MONGKUT'S UNIVERSITY OF TECHNOLOGY NORTH BANGKOK, THAILAND
 FERN UNIVERSITY IN HAGEN, GERMANY
 OKLAHOMA STATE UNIVERSITY, USA
 CHEMNITZ UNIVERSITY OF TECHNOLOGY, GERMANY
 EDITH COWAN UNIVERSITY, AUSTRALIA
 NATIONAL INSTITUTE OF DEVELOPMENT ADMINISTRATION, THAILAND
 RAJAMANGALA UNIVERSITY OF TECHNOLOGY THANYABURI, THAILAND
 RAJAMANGALA UNIVERSITY OF TECHNOLOGY KRUNGTHEP, THAILAND
 RAJAMANGALA UNIVERSITY OF TECHNOLOGY LANNA, THAILAND

HANOI NATIONAL UNIVERSITY OF EDUCATION, VIETNAM
 MAHASARAKHAM UNIVERSITY, THAILAND
 KANCHANABURI RAJABHAT UNIVERSITY, THAILAND
 NAKHON PATHOM RAJABHAT UNIVERSITY, THAILAND
 UBON RATCHATHANI UNIVERSITY, THAILAND
 RAJABHAT MAHA SARAKHAM UNIVERSITY, THAILAND
 PRINCE OF SONGKLA UNIVERSITY, THAILAND
 COUNCIL OF IT DEANS OF THAILAND, THAILAND

ISBN 978-616-7701-85-1



9 786167 701851 >